

Chapitre 3. Généralisation du diagramme de Lexis

3.1. Introduction

Le chapitre précédent nous a permis de généraliser l'emploi du diagramme de Lexis en cas de variables temporelles. La généralisation va être maintenant poussée plus loin par abandon des variables temporelles. La démarche adoptée dans le premier point de ce chapitre (Construction de diagrammes sans variable temporelle) se veut essentiellement théorique : est-il simplement possible d'obtenir des diagramme de Lexis au départ de variables non temporelles, et ce, éventuellement sans souci de savoir si de telles constructions pourraient ou non avoir un intérêt pratique ? Ce point représente un enjeu de première importance par rapport aux objectifs de ce travail. En effet, si des données peuvent se distribuer sur un diagramme de Lexis, l'application des méthodes de l'analyse démographique en devient envisageable, du moins sur un plan théorique.

Si le premier point du chapitre se veut essentiellement théorique, le deuxième aura au contraire des visées pratiques. Nous y produirons des diagrammes de Lexis avec localisation de données, mais sans variable temporelle (des diagrammes en poids et en taille pour des données concernant des nourrissons). Par ailleurs, un exemple de diagramme basé sur les valeurs de l'espérance de vie sera également présenté. Il nous permettra d'illustrer le fait qu'un certain nombre de graphiques sont des diagrammes de Lexis qui s'ignorent.

3.2. Construction de diagrammes de Lexis sans variable temporelle

3.2.1. Variation temporelle d'une variable non temporelle

Dans le diagramme de Lexis classique, les variables en jeu sont temporelles ; les variations en cause sont également temporelles. Dans ce point, nous allons progressivement passer d'un diagramme exprimant des variations temporelles de variables temporelles à des figures proposant d'étudier la variation non temporelle de variables non temporelles. Le premier point abordera cette problématique par le biais de la variation temporelle de variables non temporelles, comme, par exemple, la taille ou le poids, variables dont la valeur se modifie dans le temps, notamment durant la première partie de la vie. Le point suivant servira à s'affranchir de la contrainte « variation temporelle », même si cette voie ne nous semble pas, *a priori*, très intéressante pour la suite du travail. Finalement, un dernier point, à notre avis également très peu porteur *a priori* pour la suite, montrera la possibilité d'établir des trajectoires forcées de pente non unitaire. Suite à ces différentes étapes, la généralisation du diagramme de Lexis aura atteint un très haut degré, ce qui est notre objectif.

Pour rappel, le diagramme de Lexis classique réduit le volume initial défini par les axes des trois variables en un plan et ce, sans perte d'information. Par ailleurs, ce diagramme comporte

des trajectoires forcées pour les unités d'observation. Les trois variables campant sur les axes du volume initial (avant réduction) se doivent de respecter des conditions très précises pour obtenir une construction plane ayant les caractéristiques voulues :

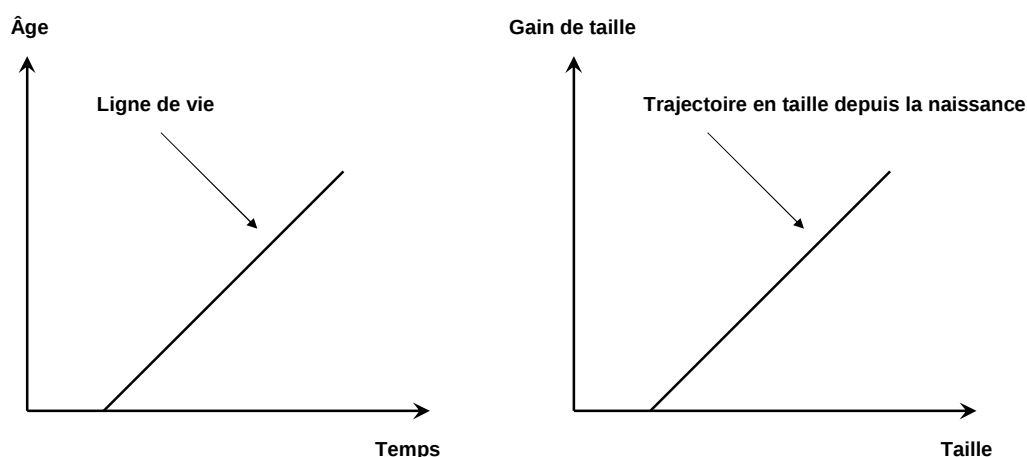
- la relation entre elles doit correspondre à une combinaison linéaire, ce qui autorise la réduction du volume en plan ;
- une des trois doit rester constante individu par individu, ce qui assure la présence de trajectoires forcées.

Les trois variables « taille », « taille à la naissance »¹ et « gain de taille depuis la naissance » satisfont pleinement les deux conditions :

- elles forment une combinaison linéaire : gain de taille = taille – taille à la naissance ;
- la taille à la naissance est constante pour un individu donné.

Des observations classées selon ces trois variables peuvent sans difficulté aucune se localiser sur un diagramme de Lexis (cf. figure 3.1). L'axe des abscisses y indiquera la taille (en remplacement du temps) ; celui des ordonnées, le gain de taille (en remplacement de l'âge) et finalement, la taille à la naissance tiendra le rôle du moment de naissance, en parasitant l'axe de la taille². Pour chaque individu, il est possible de tracer une trajectoire de taille, figurant l'évolution du gain de taille depuis la naissance en fonction de la taille en lieu et place de la classique évolution de l'âge en fonction du temps.

Figure 3.1. Diagrammes de Lexis en temps et en taille



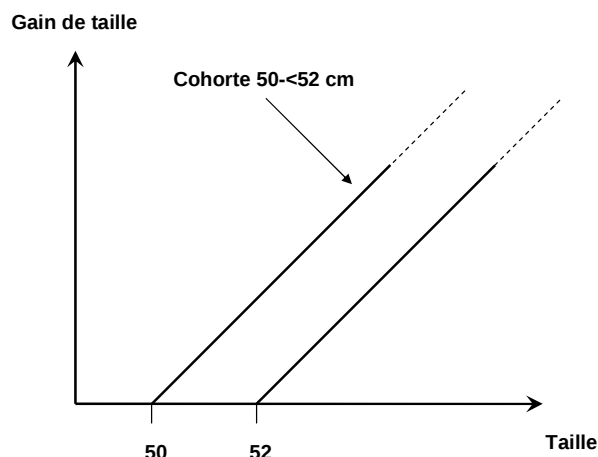
Vu les relations entre les trois variables de taille et l'identité des unités sur les deux axes, les trajectoires se présenteront toutes sous forme d'obliques croissantes inclinées à 45 degrés. La connaissance d'un seul point de la trajectoire est suffisante pour la déterminer dans son

¹ L'on pourrait s'étonner de trouver une référence temporelle explicite dans le nom de cette variable. En fait, nos mesures s'opèrent très généralement dans un contexte temporel précis. Ainsi, le temps est d'une certaine façon bien présent dans le contexte des trois variables choisies : la taille et le gain de taille depuis la naissance se mesurent à un moment donné ; la taille à la naissance se réfère à un moment bien précis. Cette constatation ne transforme pas pour autant les trois variables en variables temporelles. Elles s'expriment bien en unités de longueurs (centimètres) et nullement de temps (années, jours, secondes...) comme il sied aux variables temporelles.

² Pour les graphiques, nous avons toujours opté pour les substitutions d'axes donnant le résultat le plus immédiatement comparable à un diagramme de Lexis classique. Un diagramme en taille peut se décliner sous trois – ou six – visages, exactement comme le diagramme en temps. Il ne nous a pas paru indispensable de développer les versions concurrentes, d'autant que pour ce faire, il suffirait de suivre la logique qui a été exposée au point 2.4 du précédent chapitre à propos de la construction du diagramme de Lexis.

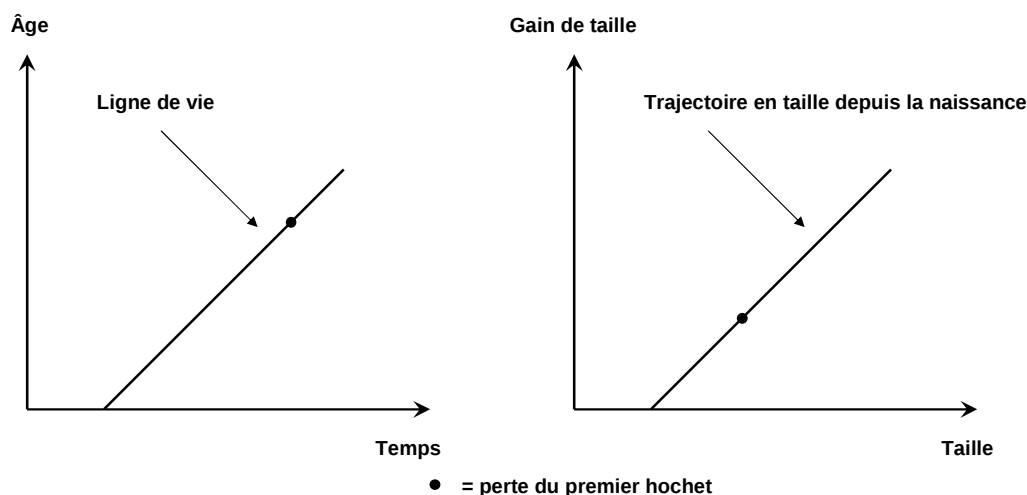
ensemble. À tous égards, ces trajectoires « forcées » jouissent des mêmes propriétés que les classiques lignes de vie. Et les cohortes dans tout cela ? Une cohorte regroupera tous les individus partageant une même gamme de taille à la naissance, par exemple entre 50 et 52 centimètres exacts³ (50->52 cm). Elle sera délimitée par les trajectoires forcées correspondant aux tailles à la naissance de 50 et 52 centimètres exacts (cf. figure 3.2).

Figure 3.2. Diagramme de Lexis et cohorte en taille



Autant il est possible de localiser des événements sur les lignes de vie d'un diagramme de Lexis classique, autant il est possible de le faire sur un diagramme de Lexis en taille. Quels événements ? Tous, aussi bien les événements démographiques (le décès, l'émigration...) que non démographiques (le doublement du poids depuis la naissance, le premier anniversaire, le début de la marche, la perte du premier hochet...). Ainsi, sur la figure 3.3, la perte du premier hochet a été localisée sur les trajectoires d'âge et de taille d'un individu ; l'événement localisé pourrait tout aussi bien être la contraction d'une maladie qu'une migration ou le doublement de poids depuis la naissance.

Figure 3.3. Localisation d'un événement sur des diagrammes en temps et en taille



³ La taille à la naissance est théoriquement une variable continue. Dans la pratique, elle s'exprime en centimètres. La cohorte en cause regroupera donc les individus dont la taille déclarée à la naissance était soit de 50, soit de 51 centimètres. La gamme de valeurs effectivement couverte par la cohorte dépendra de la façon dont la valeur est approximée, par arrondi ou troncature.

Il est à noter que la localisation sur le diagramme en taille se fait indépendamment du temps. Pour un individu, un événement peut correspondre à un âge et pour un autre individu, à un âge tout à fait différent. De même, sur le diagramme en temps, la localisation se fait en gommant complètement la notion de taille.

Le tableau 3.1 reprend différents exemples de variables qui satisfont aux deux conditions. La première colonne reprend l'exemple classique (temps, moment de naissance et âge) ; il est ainsi aisé de comparer les différents exemples. Le cas 2 reprend l'exemple du diagramme en taille. Pour les cas 3, 4 et 5, il est évidemment possible de produire un diagramme de Lexis et de constituer des cohortes respectivement de poids, de chiffre d'affaire et d'effectif de personnel.

Tableau 3.1. Différents exemples de variables

Cas 1	Cas 2	Cas 3	Cas 4	Cas 5
Temps	Taille	Poids	Chiffre d'affaire	Nombre d'employés
Moment de naissance	Taille à la naissance	Poids à la naissance	Chiffre d'affaire à un événement-origine	Nombre d'employés à un événement-origine
Age	Gain de taille depuis la naissance	Gain de poids depuis la naissance	Gain du chiffre d'affaire depuis l'évènement-origine	Gain dans le nombre d'employés depuis l'évènement-origine

En fait tous ces exemples sont semblables. Même si ce n'est pas habituel dans le langage courant, le moment de naissance correspond au « temps à la naissance », la naissance étant l'évènement-origine ; l'âge correspond, quant à lui, à « un gain de temps depuis la naissance ». Chacun des cas se compose donc des mêmes trois éléments :

- une variable « absolue » de référence (le temps, la taille...), expression généralisant la durée de référence utilisée dans le point 2.5.3. Généralisation du concept de « cohorte » ;
- la valeur de cette variable absolue à un événement-origine ;
- la variation de cette variable absolue depuis l'évènement-origine.

Ensemble, ces trois variables constituent le cadre de référence dans lequel l'analyse se déroule. Même si cela peut paraître évident, il nous semble utile d'insister sur le fait que la variable absolue doit remplir certaines caractéristiques pour pouvoir jouer son rôle :

- elle doit être susceptible de prendre, individu par individu, des valeurs différentes au moment de l'évènement-origine ; dans le cas contraire, il ne sera pas possible de définir plusieurs cohortes puisque sa valeur à l'évènement-origine sera la même pour toutes les unités d'observation ;
- elle doit être susceptible de varier dans le temps ; dans le cas contraire, il ne sera pas possible d'utiliser la troisième coordonnée, équivalente à l'âge.

À propos de l'évènement-origine, dans les trois premiers cas, il s'agit du même, à savoir la naissance. Par contre dans les deux autres, cet événement-origine n'est pas précisé. Il pourrait tout aussi bien s'agir d'une date précise (correspondant à l'entrée en vigueur d'une nouvelle loi économique ou sociale ou à un crac boursier, par exemple) ou alors d'un moment variable d'une firme à l'autre correspondant, par exemple, au moment où la firme commence à exporter ; dans ce dernier cas, il s'agirait donc d'une référence temporelle « flottante » par rapport au temps-calendrier.

Par ailleurs, si l'on veut quitter le point de vue purement théorique pour atteindre éventuellement une certaine portée pratique, l'évènement-origine devrait correspondre à un phénomène au moins potentiellement significatif pour les unités d'observation et leur

« évolution » par rapport à la variable absolue, comme le sont éventuellement la naissance ou l'entrée en vigueur d'une loi ou le début des exportations.

Si, sur le plan théorique, nous venons de montrer que des variables non temporelles pouvaient remplacer les variables classiques des démographes, il n'en faut pas moins souligner des différences. Dans le diagramme de Lexis classique, le temps, ainsi que les lignes de vie et l'âge, présentent un déroulement immuable, linéaire, et continu : ils progressent de manière régulière et inexorable⁴. Qu'en est-il pour d'autres types de trajectoires ?

Au contraire du temps, le poids ne varie pas de manière continue et inexorable. Un individu peut voir stagner son poids pendant un certain temps, puis se modifier à des rythmes différents selon les périodes. Par ailleurs, si la taille est un phénomène irréversible, du moins durant une bonne part de la vie et en dehors de situations exceptionnelles, le poids, par contre, est réversible : un individu peut perdre du poids, comme c'est notamment le cas après la naissance ; il faut donc envisager des variations négatives.

Cela ne remet pas en cause le caractère forcé des trajectoires : si, vu son évolution, une unité d'observation se déplace, elle ne peut le faire qu'en suivant sa trajectoire, avec éventuellement des allers et retours. De même, la localisation d'événements sur ces trajectoires reste licite sur le plan théorique et potentiellement significative sur le plan pratique. Nous reviendrons sur ce sujet au point 3.2.4.

3.2.2. Variation non temporelle d'une variable non temporelle

Ce point vise à étendre l'utilisation du diagramme de Lexis au cas de la variation non temporelle d'une variable. L'exemple proposé n'offrira d'autre intérêt que celui d'illustrer cette possibilité sur un plan purement théorique, sans chercher nullement à atteindre une quelconque portée pratique.

Dans une masse d'air, la température diminue généralement avec l'altitude⁵. Le gradient de température (variation de la température exprimée en degrés Celsius par 100 mètres) varie selon les circonstances de 0,3°C à 1,5°C par 100 mètres. Par ailleurs, dans certaines conditions particulières, une inversion de température peut se mettre en place : à partir d'une certaine altitude, la température augmente.

Pour les besoins d'une recherche à propos du gradient de température, supposons une série d'endroits sélectionnés à la surface du globe. Ces endroits se situent à différentes latitudes, longitudes et altitudes. Pour chacun d'eux, la température a été enregistrée simultanément au sol et à 1 000 mètres au-dessus du sol (ci-après dénommée « température à 1 000 mètres »). Ces données permettent évidemment de calculer une variation de température (« ΔT° ») :

$$\Delta T^\circ = T^\circ \text{ à 1 000 mètres} - T^\circ \text{ au sol.}$$

Ainsi définie, la variation peut être aussi bien positive que négative. Supposons que la température ait aussi été mesurée à 2 000 mètres au-dessus du sol (toujours en respectant la simultanéité des relevés de température pour un même point géographique). La variation par

⁴ On reviendra plus loin sur ces caractéristiques du temps (cf. point 3.2.5 et annexe 3.1).

⁵ À ce sujet, cf. notamment B. Lepettre (1995).

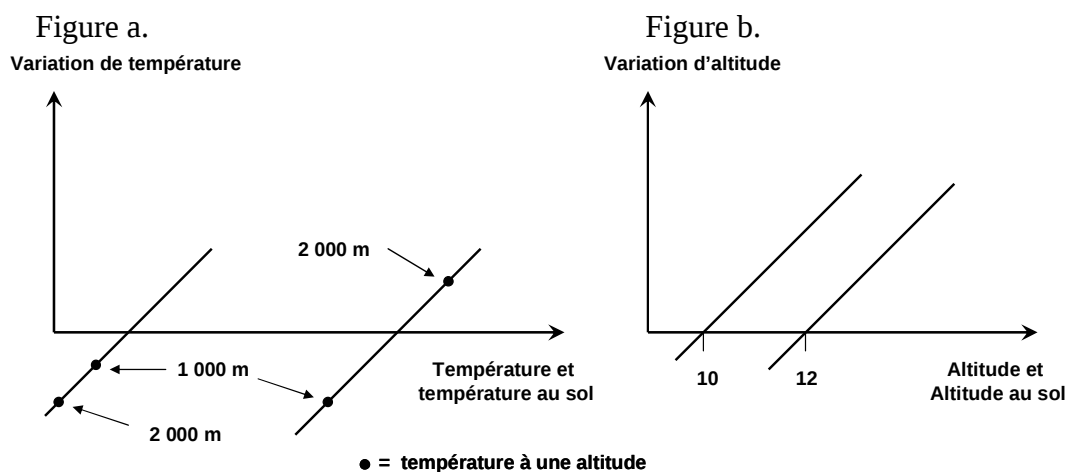
rapport au sol pourra aussi se calculer, la base de référence – soit la température au sol – étant immuable pour un endroit, vu la simultanéité des mesures.

Les trois variables « température », « température au sol » et « variation de température par rapport au sol » (qui forment, dans le cas présent, le cadre de référence de l'étude) respectent les deux conditions à l'établissement d'un diagramme de Lexis :

- elles se combinent dans une relation additive : $\Delta T^\circ = T^\circ - T^\circ \text{ au sol}$;
- pour une unité d'observation (soit un endroit à la surface du globe), une variable a une valeur constante : la température au sol.

En conséquence, des observations réalisées dans un tel cadre de référence peuvent se disposer sur un diagramme de Lexis (cf. figure 3.4.a). L'axe des abscisses y supporte à la fois la température et la température au sol et celui des ordonnées, la variation de température par rapport au sol. Par référence à la terminologie introduite sous le tableau 3.1, la température joue le rôle de variable absolue ; la température au sol, celui de la mesure de cette variable absolue à l'évènement-origine que constitue le sol et la variation de température, celui de la variation de la variable absolue par rapport à l'évènement-origine.

Figure 3.4. Variation non temporelle et diagramme de Lexis



Dans une telle figure, un point sur une trajectoire peut représenter, par exemple, une altitude au dessus du sol par rapport à un endroit géographique, altitude pour laquelle on a relevé la température. Une altitude (c'est-à-dire un évènement⁶) se localise en fonction de ces coordonnées : température, température au sol et/ou variation de température depuis le sol. Par ailleurs, quelle que soit l'altitude par rapport au sol, tous les points attachés à un endroit géographique ne peuvent que se distribuer le long d'une oblique croissante de pente unitaire et interceptant l'axe des abscisses à la valeur de la température au sol. Ces obliques constituent donc des trajectoires forcées.

Sur cette figure 3.4.a, les trajectoires forcées se développent aussi bien au-dessus qu'au-dessous de l'axe des abscisses ; en effet, les variations peuvent aussi bien être négatives que positives. La trajectoire de gauche illustre le cas d'un endroit surmonté d'une masse d'air où la température diminue avec l'altitude (température plus basse à 2 000 mètres qu'à 1 000 ou à 1 000 qu'au sol) et la deuxième, celui d'une inversion de température : à 2 000 mètres, la température est plus élevée qu'à 1 000 ou même qu'au sol.

⁶ Sur toutes ces figures, un point indique la localisation d'un évènement ; à ce sujet, cf. le point 3.2.4. Évènements et coordonnées : généralisation.

Comme le montrent ces exemples, une unité d'observation peut opérer des allers et retours sur sa trajectoire, mais sans jamais pouvoir la quitter. Par ailleurs, un même point d'une trajectoire peut supporter plusieurs événements : ce serait le cas si, suite à une inversion de température, les températures au sol et à 2 000 mètres étaient identiques et supérieures à celle mesurée à 1 000 mètres. Il s'agit là d'une différence par rapport au diagramme de Lexis classique, situation déjà rencontrée au point précédent avec le diagramme de Lexis en poids par exemple.

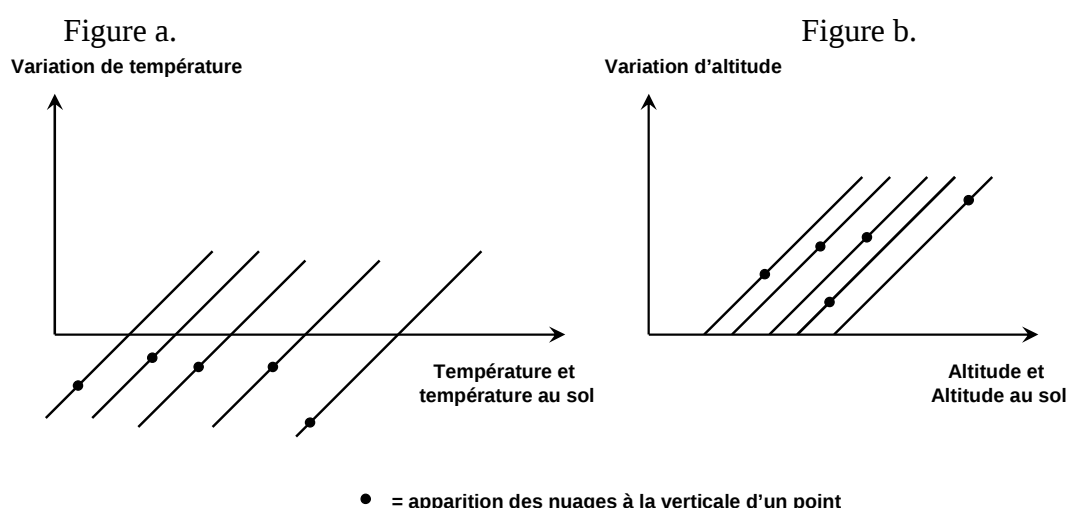
Comme le suggère la figure 3.4.b, une cohorte regroupera l'ensemble des endroits du sol partageant une même gamme de températures (par exemple entre 10 et 12°C).

Mis à part l'altitude, d'autres phénomènes pourraient aussi se représenter sur un tel diagramme de Lexis, par exemple, l'apparition des nuages. Comment se forment les nuages⁷ ? L'air contient généralement une certaine quantité de vapeur d'eau. Selon sa température, une masse d'air ne peut dépasser une certaine teneur en vapeur d'eau ; cette quantité dite « maximale » vaut, par exemple, 9,36 et 17,15 grammes par mètre cube à respectivement 10°C et 20°C. Quand la quantité maximale est atteinte et dépassée, la condensation d'une partie de la vapeur d'eau en gouttelettes forme des nuages de façon à ajuster la quantité de vapeur résiduelle à la quantité maximale.

Supposons une masse d'air homogène verticalement en ce qui concerne l'humidité qu'elle contient (soit 9,36 grammes par mètre cube) et dont la température au sol est de 20°C. Si son gradient de température est de 1°C par 100 mètres, à 1 000 mètres, la température sera de 10°C. Si l'on s'élève encore, la température, en diminuant, fera passer la quantité maximale sous les 9,36 grammes par mètre cube. En conséquence, la condensation se produira. Les nuages commenceront donc à se former à 1 000 mètres et à la température de 10°C.

Selon les conditions de température au sol ainsi que selon l'évolution verticale de la température et celle de l'humidité (supposée constante par hypothèse), le point à partir duquel les nuages se formeront peut varier d'un endroit à l'autre. Ces points pourraient se disposer sur un diagramme de Lexis (*cf.* figure 3.5.a).

Figure 3.5. Variation non temporelle et diagramme de Lexis : deux cas



Le même phénomène pourrait donner lieu à l'établissement d'un autre diagramme de Lexis en altitude cette fois (*cf.* figure 3.5.b) : l'altitude et l'altitude au sol se disposeraient sur l'axe des

⁷ À ce sujet, *cf.* notamment B. Lepettre (1995).

abscisses en lieu et place de la température et de la température au sol ; la variation d'altitude relayerait la variation de température sur l'axe vertical. Chaque endroit – chaque altitude au sol et plus chaque température au sol – se matérialiserait par une oblique croissante de pente unitaire.

Par ailleurs, ces exemples respectent le schéma qui a été proposé sous le tableau 3.1, à savoir, qu'on y retrouve :

- une variable absolue, soit la température ou l'altitude ;
- la valeur de cette variable absolue à un événement-origine, soit la température ou l'altitude au sol ;
- la variation de cette variable absolue par rapport à l'évènement-origine, soit la variation de température par rapport au sol ou la variation d'altitude par rapport au sol.

Bref, de telles observations peuvent très bien s'agencer sur une figure ayant toutes les caractéristiques du diagramme de Lexis, faisant apparaître des trajectoires forcées et des cohortes. Il s'agit donc bien de diagrammes de Lexis sans variation temporelle, puisque les différentes mesures y intervenant se font simultanément.

Ce type de diagramme nous permet donc de passer à une figure construite non plus en fonction de la variation temporelle d'une variable de nature quelconque (point précédent), mais bien en fonction de la variation de nature quelconque d'une variable de nature quelconque. De la sorte, la généralisation atteint un degré élevé, du moins sur le plan théorique, ce qui était notre objectif, et indépendamment donc de toute portée pratique immédiate.

3.2.3. Diagramme de Lexis à obliques de pente quelconque

Comme dans le point précédent, et même bien plus encore, l'exemple utilisé ici ne cherche pas à montrer la portée pratique de l'élaboration d'un diagramme de Lexis, mais vise à illustrer la possibilité théorique d'obtenir des diagrammes de Lexis présentant une caractéristique particulière, à savoir des obliques à pente quelconque.

Soit une entreprise bénéficiant chaque année d'une aide publique (« A »), qui s'ajoute à son bénéfice (« B »). Pour un de ses employés, le salaire (« S ») correspond à un quart du bénéfice augmenté d'un tiers de l'aide publique :

$$S = 1/4 B + 1/3 A$$

Les trois variables en jeu remplissent bien les conditions pour opérer la réduction d'axes. Décidons de porter le bénéfice en abscisse, de supprimer l'axe de l'aide et de reporter cette variable sur l'axe des ordonnées (*cf.* figure 3.6). Sur un tel diagramme, une valeur donnée de la variable « aide » (soit *k*) se traduira par une oblique d'équation :

$$S = 1/4 B + 1/3 k$$

La pente de cette droite vaut 1/4 ; les valeurs à mettre sous celles du bénéfice pour créer l'échelle de l'aide se calculent en annulant le salaire :

$$S = 1/4 B + 1/3 A$$

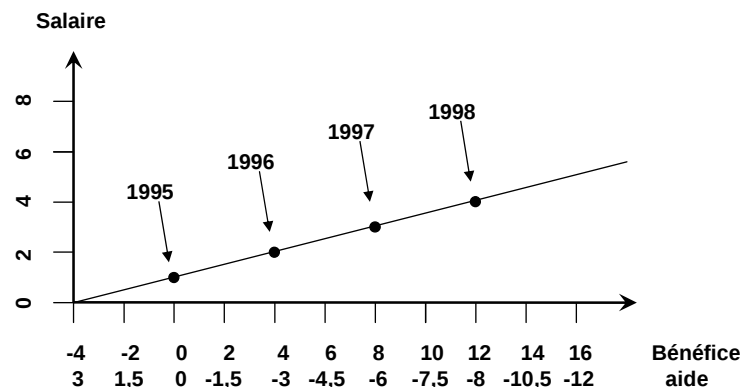
$$\Rightarrow 0 = 1/4 B + 1/3 k$$

$$\Rightarrow k = -3/4 B$$

Par exemple, si le bénéfice vaut 10, l'aide vaut -7,5 ; si le bénéfice vaut -4, l'aide vaut 3 (cf. figure 3.6). Supposons que :

- la façon de calculer le salaire de l'individu en cause reste la même d'une année à l'autre ;
- l'aide reçue par l'entreprise s'élève constamment à 3 ;
- le bénéfice de l'entreprise a varié comme suit : 1995 : 0 ; 1996 : 4 ; 1997 : 8 ; 1998 : 12.

Figure 3.6. Diagramme de Lexis à obliques de pente non unitaire



En conséquence, le salaire de l'individu s'est élevé pour ces différentes années respectivement à 1, 2, 3 et 4. Les points localisant les différentes années ne pourront s'écarter de la diagonale matérialisant l'aide perçue par l'entreprise, soit celle interceptant l'axe des abscisses à la valeur -4 en termes de bénéfice et 3 en termes d'aide (cf. figure 3.6). Pour s'en écarter, il faudrait que le montant de l'aide se modifie, possibilité exclue par hypothèse. La diagonale est donc bien une trajectoire forcée, du moins tant que l'aide reste stable.

Supposons maintenant que :

- la façon de calculer le salaire de l'individu en cause reste la même d'une année à l'autre ;
- le salaire perçu par l'individu doit constamment rester à 3 ;
- l'aide perçue par l'entreprise varie d'une année à l'autre de façon à assurer la constance du salaire en dépit des fluctuations du bénéfice.

Si les bénéfices ont fluctué comme indiqué plus haut, la valeur de l'aide s'est élevée successivement à 9 en 1995 ; 6 en 1996 ; 3 en 1997 et 0 en 1998. Décidons, pour créer un diagramme de Lexis avec ces données, de supprimer l'axe du salaire et de reporter cette variable sur l'axe du bénéfice mis en abscisse, l'aide étant sur l'axe vertical. Le salaire s'y matérialisera par une oblique décroissante :

$$S = 1/4 B + 1/3 A, \text{ avec } S = 3$$

$$\Rightarrow A = 9 - 3/4 B.$$

Sa pente sera donc de -3/4. Ensuite, en annulant l'aide, on détermine que les valeurs à mettre sous l'échelle du bénéfice pour créer celle du salaire se calculent en prenant un quart du montant du bénéfice (cf. figure 3.7). Pour l'individu en cause, les points représentant les

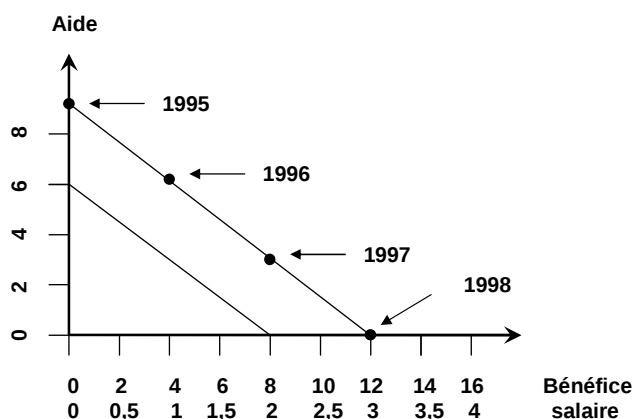
années ne pourront s'écarter d'une trajectoire forcée passant par les points de coordonnées (bénéfice = 0 ; aide = 9) et (bénéfice = 12 ; aide = 0).

Si, pour un employé d'une autre firme suivant les mêmes règles pour le calcul des rétributions, le salaire doit toujours valoir 2, les points ne pourront jamais s'écarter d'une oblique décroissante d'équation :

$$A = 6 - 3/4 B.$$

Sa pente sera donc aussi de $-3/4$. Cette oblique sera parallèle à la précédente et interceptera l'axe des abscisses à la valeur 8 en termes de bénéfice et 2 en termes de salaire (cf. figure 3.7). Comment définir dans ce contexte des cohortes ? Une cohorte regroupera l'ensemble des individus qui partagent la même gamme de salaire. Ainsi, sur la figure 3.7, les deux diagonales délimitent la cohorte des individus disposant d'un salaire compris entre 2 et 3. Ne pourront entrer dans cette cohorte que les individus dont le salaire est compris entre 2 et 3 et correspond à un quart du bénéfice augmenté d'un tiers de l'aide, ce qui limite de beaucoup la portée pratique potentielle de la figure, mais ce n'était pas l'objectif poursuivi.

Figure 3.7. Diagramme de Lexis à obliques de pente non unitaire. Deuxième cas



Les deux dernières figures montrent des exemples de diagrammes de Lexis à obliques de pente non unitaire ; du point de vue purement théorique, ces figures sont bien des diagrammes de Lexis à part entière, avec possibilité d'identifier des trajectoires forcées et des cohortes. En fait, ces figures nous ramènent à l'exemple purement mathématique exposé au point 2.4. Conditions mathématiques de l'élaboration d'un diagramme de Lexis. À notre sens, il s'agit de la généralisation la plus haute qui puisse être proposée pour l'élaboration de diagrammes en plan permettant le recours à trois coordonnées et comportant des trajectoires forcées.

Il est à noter ici que les éléments en jeu ne suivent pas le schéma proposé sous le tableau 3.1 : aucune variable ne peut se qualifier d'absolue et donc la valeur de cette variable à un événement-origine ou sa variation ne peuvent trouver de correspondance. Nonobstant cette remarque, les figures qui précèdent présentent bien toutes les caractéristiques d'un diagramme de Lexis.

Il est à rappeler que ce qui vient d'être développé dans ce point 3.2.3 ne poursuivait qu'un but purement théorique de généralisation de l'élaboration du diagramme de Lexis. Sans réelle portée pratique exploitable, cet aspect (pente non unitaire) est à considérer comme une parenthèse purement théorique et le cas envisagé sera abandonné dans la suite de l'exposé, pour laquelle la généralisation atteinte en fin de point 3.2.2 est largement suffisante.

3.2.4. Évènements et coordonnées : généralisation

Au cours de la généralisation proposée dans les points précédents (et sans prendre en compte le point 3.2.3), certains concepts ont été employés en dehors de leur contexte habituel. Ainsi, la notion d'évènement a été utilisée pour désigner des éléments aussi disparates que la perte du premier hochet, l'altitude à la verticale d'un point ou la formation des nuages. Un évènement est ici compris comme un fait de nature quelconque affectant une unité d'observation. Un évènement se définit donc de manière moins stricte qu'en situation classique. Pour Allison, par exemple :

« an event consists of some qualitative change that occurs at a specific point in time. One would not ordinarily use the term 'event' to describe a gradual change in some quantitative variable. The change must consist of a relatively sharp disjunction between what precedes and what follows »⁸.

Dans cette définition, trois points ressortent :

- le changement subi doit être qualitatif ;
- il doit être synonyme d'une disjonction tranchante entre ce qui précède et ce qui suit ;
- pour un changement graduel mesuré à propos d'une variable quantitative, on devrait éviter de recourir au vocable d'évènement.

Le premier mariage est un bon exemple d'évènement suivant cette définition : l'individu change de statut en quittant l'état de célibataire pour accéder à celui de personne mariée, ce qui marque une rupture sur le plan social. Dans l'exemple de la figure 3.4, les points représentent des altitudes. « 1 000 mètres au-dessus du sol » ne marque pas une rupture qualitative pour une unité d'observation, soit un point géographique ; de plus, une altitude est un point particulier dans le changement graduel d'une variable franchement quantitative ! Bref, les critères identifiés au départ de la définition d'Allison ne sont pas respectés. Serait-ce une raison suffisante pour retirer la qualité d'évènement à ce fait ?

À notre avis, non. L'objectif du diagramme de Lexis et des analyses qui y font généralement suite est l'étude de la survenance d'un fait affectant des unités d'observation et se déroulant dans un cadre de référence précis. Le tableau 3.2 reprend six exemples d'évènements en identifiant à chaque fois le fait étudié, les unités d'observation concernées et le cadre de référence défini par les trois variables le constituant.

Dans chacun des cas, la localisation des faits ou évènements sur un diagramme de Lexis approprié s'opère selon les mêmes principes que sur un diagramme classique. Les faits en cause se matérialisent tous par des points sur les trajectoires forcées définies dans le cadre de référence choisi. Il n'y a aucune raison *a priori* de distinguer de ce point de vue un cas des autres, à tout le moins sur le plan théorique. Par ailleurs, ces exemples peuvent déboucher sur des questions à notre avis équivalentes et toutes aussi licites les unes que les autres, toujours sur un plan théorique :

- cas 1 : quel est l'âge moyen au 1^{er} mariage ?
- cas 2 : quel est l'âge moyen à la perte du 1^{er} hochet ?
- cas 3 : quel est le gain de taille moyen à la perte du 1^{er} hochet ?

⁸ P. Allison (1984), p. 9. À ce sujet, cf. aussi C. Vandeschrick (2004), p. 195.

- cas 4 : quelle est la variation moyenne de température à 1 000 mètres du sol ?
- cas 5 : quelle est la variation moyenne de température par rapport au sol à la formation des nuages ?
- cas 6 : quelle est la variation moyenne d'altitude par rapport au sol à la formation des nuages ?

Tableau 3.2. Évènements : généralisation

	Fait, évènement	Unités d'observation	Cadre de référence
Cas 1 (cas classique)	1 ^{er} mariage	individus	- temps - moment de naissance - âge
Cas 2 (cf. figure 3.1)	perte du 1 ^{er} hochet	individus	- temps - moment de naissance - âge
Cas 3 (cf. figure 3.2)	perte du 1 ^{er} hochet	individus	- taille - taille à la naissance - gain de taille
Cas 4 (cf. figure 3.4)	1 000 m. au-dessus du sol	points géographiques	- température (t°) - température au sol - variation de t° par rapport au sol
Cas 5 (cf. figure 3.5.a)	apparition des nuages	points géographiques	- température (t°) - température au sol - variation de t° par rapport au sol
Cas 6 (cf. figure 3.5.b)	apparition des nuages	points géographiques	- altitude - altitude au sol - variation d'altitude par rap. au sol

Ces exemples montrent, à notre avis, qu'un évènement étudié peut être de nature quelconque et ne pas supposer une rupture pour l'individu sous observation. Notons toutefois que la perte du 1^{er} hochet peut, à la limite, être considérée comme une rupture autant que le 1^{er} mariage, en faisant passer l'individu du statut de disposant encore de son 1^{er} hochet à celui de n'en disposant plus ! La même remarque peut aussi valoir pour deux autres faits du tableau 3.2 :

- l'altitude de 1 000 mètres fait quitter à l'unité d'observation sa 1^{re} tranche de 1 000 mètres pour la faire entrer dans la deuxième⁹ ;
- l'apparition des nuages marque le passage de la non-saturation de l'air en vapeur d'eau à sa saturation.

Ces ruptures sont certes inégalement spectaculaires ; toutefois, rien n'empêche, sur le plan théorique de nouveau, de les considérer en tant que rupture au même titre qu'un 1^{er} mariage ou une migration. Pour la suite de ce texte, nous garderons donc une définition la plus large possible pour le concept d'évènement. Nous proposons la définition suivante : « circonstance pour laquelle les coordonnées formant le cadre de référence choisi ont été enregistrées pour les unités d'observation » ; si un fait répond à cette définition, il est localisable sur la trajectoire de l'individu et sa survenance peut être analysée dans le cadre de référence choisi.

⁹ Le 1^{er} anniversaire fait quitter à l'individu sa 1^{re} année de vie pour l'introduire dans la 2^e. Cet exemple est en fait très proche de l'altitude de 1 000 mètres ; dans les deux cas, il s'agit d'une valeur particulière dans le changement progressif d'une variable quantitative. Cela n'empêchera nullement les démographes de calculer, par exemple, la probabilité d'atteindre le 1^{er} anniversaire, soit de survivre de la naissance au 1^{er} anniversaire !

Cette définition large, voire lâche, du concept d'évènement offre l'avantage d'inclure sa conception classique tout en laissant une grande latitude pour le choix des faits à analyser et des cadres de références pour mener ces analyses.

En abandonnant l'aspect purement théorique de la question, ajoutons toutefois que, sur le plan pratique, la pertinence d'une analyse ne pourra être assurée que dans la mesure où l'évènement étudié est potentiellement porteur de sens pour les individus observés et que le cadre de référence reprend des variables potentiellement explicatives du phénomène étudié dans la population visée. Par ailleurs, pour constituer des cohortes sur base de la proximité de la variable de référence à l'évènement-origine, il faut au moins suspecter que cette coordonnée est susceptible d'influencer le comportement des individus par rapport au phénomène étudié.

La généralisation du concept d'évènement qui vient d'être proposée doit rejaillir aussi sur certaines expressions pour éviter de retomber sous la coupe d'un réflexe exclusivement temporel et assurer ainsi la généralisation en ce qui concerne le cadre de référence. Par exemple, la localisation d'un décès sur le diagramme classique peut se faire via les deux coordonnées suivantes : la date et l'âge au moment du décès. Si la localisation devait se faire sur un diagramme en taille, ces deux coordonnées pourraient être remplacées par la taille et le gain de taille au moment du décès. Dans ces énoncés, l'expression « au moment du » nous semble superflue ; par ailleurs, elle rappelle trop l'idée de référence au temps qui est soit déjà comprise naturellement dans l'âge et le temps, soit sans réel intérêt quand on veut travailler dans un cadre de référence en taille et non plus en temps. Nous proposons donc d'adopter des énoncés simplifiés :

- date et âge au décès ;
- taille et gain de taille au décès.

En fait, ces énoncés se construisent sur deux modèles, à savoir :

- la valeur de la variable absolue à l'évènement, pour la date et la taille ;
- la variation de la variable absolue à l'évènement par rapport à sa valeur à l'évènement-origine, pour l'âge et le gain de taille.

En suivant ces modèles d'énoncés, mais en se limitant à la coordonnée indiquant la valeur de la variable absolue à l'évènement, on aura :

- la température à 1 000 mètres, dans le cas de la figure 3.4 ;
- la température à la formation des nuages, dans le cas de la figure 3.5.a ;
- l'altitude à la formation des nuages, dans le cas de la figure 3.5.b ;
- la taille au premier anniversaire, si l'évènement étudié est le premier anniversaire et la variable absolue, la taille.

Tous ces exemples suivent bien la même structure générale : valeur de la variable absolue à l'évènement ; par ailleurs, ils soulignent la nature très disparate des évènements à localiser sur un diagramme de Lexis. Le remplacement de la mention « au moment de l'évènement » par « à l'évènement » a été proposé pour s'affranchir par rapport à la tyrannie du temps, phase indispensable pour parvenir à concevoir des cadres de référence non basés sur le temps. Ces notions bien établies permettent de généraliser le diagramme de Lexis en passant de la représentation classique de la variation temporelle d'une variable temporelle à celle de la variation de type quelconque d'une variable de nature quelconque.

3.2.5. Trajectoires et trajets

Jusqu'à présent, nous avons utilisé la seule notion de trajectoire pour décrire le parcours d'un individu sur le diagramme de Lexis. Suite à la généralisation qui vient d'être proposée à son sujet et dans l'optique de la généralisation des méthodes de l'analyse démographique, il faut introduire la notion supplémentaire de trajet ; c'est l'objet de ce point. On partira de la notion de trajectoire pour définir ensuite ce qu'est un trajet.

L'exemple de la température (point 3.2.2) propose des trajectoires se développant aussi bien en dessous qu'en dessus de l'axe des abscisses. En effet, la température peut aussi bien varier à la hausse qu'à la baisse quand l'altitude augmente. En tous points, la trajectoire indique la valeur de la variation de température par rapport au sol pour une température donnée. La trajectoire se présente sous la forme d'une fonction linéaire croissante continue de pente 45° et interceptant l'axe des abscisses à la valeur de la variable absolue à l'évènement-origine.

La trajectoire se définit comme l'ensemble des points du diagramme qu'un individu peut en théorie fréquenter étant donné la valeur de la variable absolue à l'évènement-origine qui l'a fait rentrer dans l'observation. *A priori*, il n'y a pas de raison de limiter une trajectoire¹⁰ : elle peut se prolonger (indéfiniment) vers le haut ou vers le bas ; quelle que soit la valeur de la variable absolue, la trajectoire montrera toujours la variation de la variable absolue (par rapport à sa valeur à l'évènement-origine) y correspondant.

Un individu va-t-il fréquenter tous les points de sa trajectoire ? Pas forcément : en se limitant à la basse atmosphère, la température à la verticale d'un point ne dépassera pas certaines limites, ni vers le haut, ni vers le bas. Une unité d'observation n'exploitera qu'une partie de sa trajectoire. Les points fréquentés par un individu ne le seront-ils qu'une seule fois ? Pas forcément : en suivant la verticale d'un point, la température peut passer successivement par des épisodes de croissance et de décroissance de température et donc passer plusieurs fois par la même valeur. Comme le montre cet exemple, il faut distinguer la trajectoire d'un individu (ensemble des points théoriquement accessibles pour une unité d'observation) de son trajet effectif sur sa trajectoire (l'ensemble des points de la trajectoire effectivement fréquentés par une unité d'observation).

Le trajet ne peut donc se dessiner que sur la trajectoire, mais ces deux notions ne sont pas assimilables. Les caractéristiques d'une trajectoire et d'un trajet sont en rapport respectivement avec celles de la variable absolue et celles de la variation de la variable absolue par rapport à l'évènement-origine. Pour expliciter ce point, nous allons prendre trois exemples : la situation classique avec le temps comme variable absolue, la température à la verticale d'un point, le salaire dont disposent les individus.

Ces trois variables absolues (le temps, la température et le salaire) peuvent être considérées comme (implicitement) continues. En conséquence, sur un diagramme de Lexis, les trajectoires apparaîtront sous forme de fonctions linéaires continues croissantes de pente unitaire, pour autant que l'on attribue les axes des abscisses et des ordonnées respectivement à la variable absolue et à sa variation par rapport à l'évènement-origine. Le fait que la température ou le salaire peuvent aussi bien varier à la hausse qu'à la baisse par rapport à la valeur à l'évènement-origine se traduira par une trajectoire se développant aussi bien au-dessus qu'au-dessous de l'axe des abscisses¹¹, ce qui ne modifie en rien les caractéristiques de la fonction.

¹⁰ Dans le cas de la température, on évitera de descendre sous le zéro absolu, soit - 273,15 degrés centigrades.

¹¹ À la limite, en acceptant des âges négatifs, les trajectoires du cas classique partagent aussi ce caractère : pour une date antérieure à sa date de naissance, l'âge d'un individu serait négatif !

Pour obtenir une trajectoire non continue, il faudrait que la variable absolue soit elle-même non continue, comme dans le cas du nombre de travailleurs ayant signé un contrat d'emploi dans une entreprise (cf. tableau 3.1) ; dans ces circonstances, la trajectoire ne serait définie que pour les nombres entiers. Par contre, la trajectoire suivra un alignement linéaire croissant de pente unitaire, la seule caractéristique perdue étant bien la continuité. En définitive, pour autant que l'on respecte les règles d'attribution des axes, les trajectoires seront toujours de tendance linéaire, croissante de pente unitaire ; la seule différence qui puisse se marquer porte sur le caractère continu ou pas de la trajectoire. Par ailleurs, une trajectoire pourra toujours s'interpréter comme étant l'évolution de la variation de la variable absolue par rapport à sa valeur à l'évènement-origine en fonction des valeurs de la variable absolue.

Venons-en maintenant aux trajets effectivement suivis par les individus sur leurs trajectoires. Les caractéristiques des trajets sont déterminées par la façon dont évolue la variation de la variable absolue pour un individu par rapport à la valeur à l'évènement-origine. Pour régler la question du transfert des méthodes (cf. chapitre 4), il faudra déterminer les caractères uni- ou bidirectionnelle, irréversible ou pas et continu ou pas de cette variation.

Dans le cas classique, l'âge ne peut qu'augmenter après la naissance¹² ; les trajets ne peuvent comporter que des variations positives par rapport à la valeur de départ ; ils sont unidirectionnels et se développent tous uniquement au-dessus de l'axe des abscisses. Par contre, dans l'exemple du salaire, si les individus sous observation disposaient d'un salaire non nul à l'évènement-origine, la variation peut aussi bien être positive que négative ; il en va de même dans le cas de la température. Dans ces circonstances, les trajets seront donc bidirectionnels ; ils peuvent se développer aussi bien au-dessus qu'au-dessous de l'axe des abscisses.

D'une manière générale, si, les variations enregistrées par rapport à la valeur à l'évènement-origine ne sont que d'un seul signe (soit positif, soit négatif), les trajets seront unidirectionnels et bidirectionnels, si ces variations peuvent être positives ou négatives.

Le cas classique des trajets en âge est irréversible : quelqu'un qui a dépassé un âge ne pourra y revenir. L'exemple des températures et celui du salaire montrent au contraire des trajets réversibles : en s'élevant progressivement à la verticale d'un point, on peut passer par une température, la dépasser et y revenir ; pour le salaire, la situation est identique. Un trajet sera réversible si un individu sous observation peut revenir à un point de son trajet après l'avoir « dépassé » vers le haut ou vers le bas ; cela signifie donc qu'il aura connu des variations partielles de sens opposés le long de son trajet. L'irréversibilité consacre l'impossibilité de revenir plusieurs fois à une même variation par rapport à la valeur à l'évènement-origine.

Même si les notions d'irréversibilité et d'unidirectionnalité sont proches, il ne faut pas les confondre, car leurs conséquences sur le transfert des méthodes diffèrent. Le plus souvent, ces caractères vont de pair, mais sans que ce ne soit automatique. L'unidirectionnalité signifie que, par rapport à la valeur à l'évènement-origine, les individus ne peuvent connaître que des variations d'un seul signe ; si à l'intérieur de ces variations d'un seul signe, des allers-retours sont envisageables, les trajets juxtaposent les caractères unidirectionnel et réversible.

Par exemple, à la naissance, le salaire qu'un individu peut percevoir pour son travail vaut 0. À partir du moment où il travaille, il percevra un salaire qui pourra varier au gré des circonstances à la hausse ou à la baisse, mais jamais en passant dans des valeurs négatives. On pourrait résumer cette situation en définissant la réversibilité comme une espèce de « bidirectionnalité partielle », se cantonnant éventuellement dans une gamme de variations

¹² Du point de vue de la physique, les caractéristiques du temps sont discutables, mais cela nous semble sans conséquence concrète sur le travail des démographes. À ce sujet, cf. annexe 3.1.

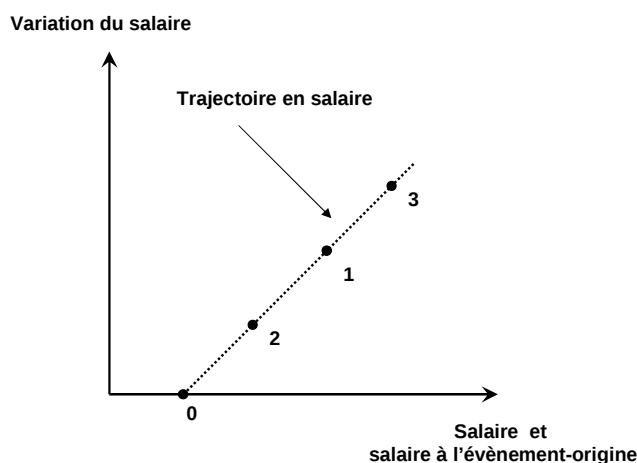
d'un seul signe par rapport à la valeur à l'évènement-origine. En théorie, un trajet peut aussi bien être réversible ou irréversible qu'il soit uni- ou bidirectionnel.

Pour avoir une chance d'être continu, un trajet doit se dessiner sur une trajectoire elle-même continue. La continuité suppose que pour passer d'une variation à une autre, un individu est obligé de passer par toutes les valeurs intermédiaires ; autrement dit, le trajet reprendra tous les points intermédiaires de la trajectoire. Le cas classique de l'âge correspond bien à cette définition : pour passer d'un âge à un autre, un individu passe par tous les âges intermédiaires. On peut aussi considérer que les changements de température à la verticale d'un point s'opèrent progressivement. Par contre, dans le cas du salaire, la variation n'est pas continue : suite à une indexation ou une augmentation du temps de travail, le salaire d'un individu passera sans transition d'une valeur à une autre, en ignorant les valeurs intermédiaires sur la trajectoire.

Quelle relation existe-t-il entre la trajectoire et le trajet ? Dans le cas classique, le trajet de l'individu terminé par son décès est simplement un extrait de la trajectoire, parcourue de manière continue et irréversible : le trajet et la trajectoire peuvent s'assimiler l'un à l'autre, du moins dans la partie de la trajectoire exploitée par le trajet. Par contre, dans le cas de la température, la possibilité des allers-retours fait qu'un trajet effectif pourra reprendre plusieurs fois un même tronçon de la trajectoire dans des sens de déplacement opposés. Le trajet n'est donc plus un simple extrait de la trajectoire ; il peut être plus compliqué tout en respectant le caractère continu.

Dans le cas du salaire, aux éventuels allers-retours s'ajoute le fait qu'un trajet sera constitué par une série de points disjoints sur la trajectoire, sans continuité. La figure 3.8 montre en pointillé la trajectoire en salaire pour un individu et les points numérotés, son trajet.

Figure 3.8. Trajectoires continues et trajets discontinus



Le point 1 indique sa position suite à un premier changement survenu après l'évènement-origine ; à cette occasion, il est passé instantanément du point 0 (sa situation à l'évènement-origine) au point 1 sans transition, en ignorant donc toutes les positions intermédiaires. Après un deuxième changement, il se retrouve à la position 2, de nouveau sans transition. Finalement, après un changement supplémentaire, il arrive en position 3. Loin d'être un extrait continu de la trajectoire, le trajet (les points 0, 1, 2 et 3) se contente d'en reprendre quelques points, avec des allers-retours. Ce trajet est unidirectionnel car ne comptant que des variations positives par rapport à l'évènement-origine, réversible vu la présence d'allers-retours et discontinu parce que comportant des bonds.

Les particularités qui viennent d'être évoquées ne remettent pas en cause la validité du diagramme de Lexis, ni son utilisation comme système de référence pour localiser des événements. Par contre, il faudra garder à l'esprit les caractéristiques des trajets au moment de passer aux calculs d'indices (probabilités ou taux) en suivant ce que les démographes ont l'habitude de faire, comme ce sera montré dans le quatrième chapitre.

3.2.6. Conclusions

Le point 3.2 nous aura permis de progresser dans l'optique de la généralisation. Par rapport au chapitre précédent, la généralisation proposée ici nous affranchit du caractère temporel des variables intervenant dans un diagramme de Lexis :

- la durée de référence se généralise via une variable absolue de nature quelconque ;
- la variation (temporelle) de la durée de référence depuis l'évènement-origine se généralise via la variation de nature quelconque de la variable absolue par rapport à sa valeur à l'évènement-origine

On passe ainsi résolument d'un diagramme montrant la variation temporelle d'une variable temporelle à la variation de nature quelconque d'une variable de nature quelconque. Pour construire un diagramme de Lexis, il suffit donc de disposer de la mesure non pas d'une même variable temporelle à au moins deux moments différents, mais bien d'une même variable en deux circonstances différentes.

Cette généralisation se reflète directement dans certaines expressions. Ainsi, la « valeur d'une variable temporelle au moment de l'évènement » devient la « valeur d'une variable à l'évènement ». Pour anodin qu'il puisse paraître, le changement dans l'expression le rend d'un emploi général, non dépendant du caractère temporel de la variable absolue et de sa variation.

Cette généralisation dans la définition du cadre de référence trouve un prolongement dans le concept d'évènement. La définition classique, qui suppose un changement qualitatif affectant les individus, peut se généraliser dans une circonstance les affectant et pour laquelle les coordonnées du cadre de référence choisi pour l'étude ont été enregistrées ; de la sorte, la localisation sur un diagramme devient possible.

Ces généralisations du cadre de référence et de la notion de concept nous auront aussi poussé à différencier la trajectoire d'un individu (ensemble des points qu'il peut fréquenter étant donné la valeur de la variable absolue à l'évènement-origine) et son trajet (ensemble des points qu'un individu a effectivement fréquentés sur sa trajectoire). Le caractère continu ou non de la trajectoire dépend du caractère continu ou pas de la variable absolue. C'est la variation de la variable absolue par rapport à l'évènement-origine qui détermine les caractéristiques du trajet :

- si elle est toujours de même signe, le trajet sera unidirectionnel, mais bidirectionnel dans le cas contraire ;
- si elle est continue, le trajet sera un extrait continu de la trajectoire ; si elle se produit par sauts, le trajet sera constitué par les points disjoints effectivement fréquentés par l'individu sur sa trajectoire ;
- si elle est réversible, le trajet pourra comporter des allers-retours sur la trajectoire.

Si les différences que nous venons de signaler ne perturbent pas l'élaboration du diagramme de Lexis en tant que telle, ni son utilisation, il sera crucial d'en tenir compte au moment d'élaborer le calcul de certains indices, comme on le verra dans le chapitre 4.

3.3. Utilisation de diagrammes de Lexis non classiques

3.3.1. Introduction

Dans le point 3.2, nous avons montré, sur un plan essentiellement théorique, la possibilité de construire des diagrammes de Lexis sans variable temporelle. Le point 3.3 aborde, pour sa part, le volet pratique de l'utilisation effective de ce type de construction pour la localisation de données réelles. Pour une question de disponibilité des données, un seul cas de cadre de référence non temporel sera envisagé ; il portera sur la variation temporelle de variables non temporelles, sans pouvoir atteindre, donc, un niveau de généralisation plus élevé qui porterait sur la variation non temporelle d'une variable non temporelle. De manière spécifique, les figures proposées porteront d'abord sur l'évolution de la taille et du poids des nourrissons depuis la naissance, exemples de variation temporelle de variables non temporelles (point 3.3.2). Ensuite, un exemple de diagramme de Lexis qui s'ignore sera proposé, ce qui nous ramènera dans le domaine des variations temporelles de variables temporelles, puisque cet exemple portera sur l'espérance de vie à différents moments (point 3.3.3).

3.3.2. Variation temporelle de variables non temporelles

3.3.2.1. Les données utilisées

Les données proviennent de la Banque de Données médico-sociale (BDMS), base gérée par l'Office de la Naissance et de l'Enfance de la Communauté française de Belgique (ONE)¹³.

Si l'ONE a mis sur pied la BDMS, c'est d'abord dans un double souci d'évaluation interne pour apprécier les résultats de ses actions et d'orientation de ses futures politiques en faveur de la petite enfance. Par ailleurs, les données de la BDMS sont mises à la disposition des chercheurs, des acteurs en matière de santé et de toute autre personne intéressée par les questions de santé de la petite enfance, moyennant une demande officielle adressée à l'ONE.

Les données sont récoltées initialement sur fiches standardisées par les Travailleurs médico-sociaux (TMS), c'est-à-dire le personnel de l'ONE qui assure le travail de suivi des grossesses et des naissances sur le terrain. La collecte peut comporter un maximum de quatre fiches pour une naissance : un « Volet prénatal » ; un « Avis de naissance » ; un « Premier contact » et un « Volet 0-1 an ». Le Volet prénatal est rempli pour les femmes qui se rendent aux consultations prénatales organisées par l'ONE ; en 1999, ce premier volet a été rempli pour 19 % des naissances¹⁴. L'Avis de naissance est complété à la maternité ; ce volet a été rempli pour 91 % des naissances en 1999. Le premier contact est proposé à toutes les mamans connues de l'ONE. En 1999, 82 % ont accepté ce contact. Finalement, sur base volontaire, les mamans peuvent bénéficier du suivi organisé par l'ONE via ses consultations de nourrissons

¹³ Pour une description de cette base de données et une analyse de contenu, cf. Office de la Naissance et de l'Enfance (2000 et 2001). Nous nous sommes largement inspiré de cette publication pour la présentation générale des données. L'extrait de la BDMS, que nous avons reçu dans le cadre d'une convention avec l'ONE, ne comporte aucune information permettant d'identifier les personnes en cause, ce qui assure le strict anonymat des données. Par ailleurs, il est à noter que le schéma de collecte décrit a été d'application jusqu'en 2002. Depuis, un nouveau schéma a été adopté.

¹⁴ Office de la Naissance et de l'Enfance (2001), p. 12.

ou, selon les besoins, de visites à domicile. Pour tous les enfants vus au moins deux fois dans ce circuit, un Volet 0-1 an est ouvert. Il sera complété soit au fur et à mesure des visites, soit à la fin de l'observation. Ce dernier volet a été rempli pour 40 % des naissances en 1999¹⁵.

Les données ainsi collectées sont rassemblées par les « Comités sub-régionaux » au nombre de six (les cinq provinces wallonnes et la région de Bruxelles-Capitale ; la province de Liège est amputée de la région germanophone). C'est à ce niveau que l'encodage est réalisé. Chaque sub-région dispose donc d'une base indépendante et intégrant les renseignements collectés sur les quatre fiches. Ces bases subrégionales sont regroupées à l'administration centrale de l'ONE pour constituer la BDMS (gestion via SPSS). La qualité des données est contrôlée une première fois lors de l'encodage et une deuxième fois à l'administration centrale de l'ONE.

Les données reçues de l'ONE portent sur les naissances des années 1994 à 2000, soit 116 309 enregistrements (liste des variables en annexe 3.2). Elles proviennent pour l'essentiel du Volet 0-1 an, éventuellement complétées par l'Avis de naissance. Le fichier était au format SPSS.

3.3.2.2. Diagramme en taille

3.3.2.2.1. Choix de l'évènement et calcul des coordonnées

Le Volet 0-1 an indique la taille des enfants à la naissance et à leur sortie de l'observation. Cette dernière se produit soit au premier anniversaire, qui correspond à la fin « normale » du suivi, soit antérieurement à cet âge par décision de la mère d'arrêter le suivi organisé par l'ONE avant l'échéance normale. L'évènement « sortie d'observation » pourra se localiser en fonction de trois coordonnées en taille de l'enfant, à savoir sa taille au moment de l'évènement et à la naissance ainsi que la variation de taille depuis la naissance. Ces trois coordonnées formeront le cadre de référence de l'analyse ; comme elles se combinent linéairement, les données à analyser pourront se localiser sur un diagramme de Lexis avec la taille en tant que variable absolue (cf. figure 3.3).

La sortie d'observation est le seul évènement - avec la naissance - pour lequel la base indique le poids et la taille¹⁶. Il n'est donc pas possible d'étudier un autre évènement que la sortie du champ d'observation (par cohorte de taille à la naissance). En fait, cette étude portant sur la sortie de l'observation est similaire à celle classique portant sur la mortalité : dans ce cas, le diagramme de Lexis sert à localiser les sorties d'observation par décès. Par ailleurs, la technique qui va être développée ici pourrait tout autant s'appliquer à d'autres évènements (comme l'acquisition d'une capacité physique, la contraction d'une maladie, la perte du premier hochet...) pour autant que la taille soit connue à ces évènements.

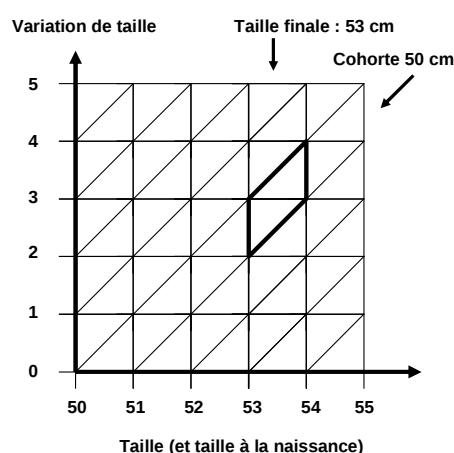
¹⁵ Le suivi des nourrissons s'opère donc sur une base volontaire. Or, il est possible qu'une sélection des observations, notamment sur base de critères socio-économiques, s'opère à cette occasion. Il est donc possible que les nourrissons dont le suivi est repris dans le fichier ne soient pas représentatifs de l'ensemble, les enfants des classes favorisées étant sans doute sous-représentés. Toutefois, vu la nature des analyses qui seront menées ci-après, il ne nous a pas semblé qu'il s'agissait là d'un point important à développer.

¹⁶ Le Dossier médical confidentiel, rempli lors de chaque consultation (hors structure ONE), offre des possibilités intéressantes. Ce dossier sert à enregistrer une série d'évènements concernant la vie de l'enfant jusqu'à l'âge de 7 ans, comme, par exemple, les changements de son régime alimentaire, l'acquisition d'une série de capacités/facultés (contrôle de la tête ; maintien de la posture debout...), etc. Vu la structure du dossier, il est possible de mettre en rapport ces évènements avec le suivi du poids, de la taille et du périmètre céphalique (également repris sur le formulaire). Il serait donc possible de cohortiser l'étude de certains évènements en fonction de ces trois variables. Nous n'avons pas eu accès à ces dossiers, dont les données ne sont par ailleurs pas encodées et/ou centralisées.

La taille à la naissance et en fin d'observation est renseignée en centimètres. Nous avons considéré que les données ont été tronquées¹⁷ et que les tailles s'exprimaient donc en centimètres révolus, par analogie avec les années révolues. Aucun effet d'attraction n'a été constaté pour cette donnée (cf. annexe 3.3). L'amplitude des cohortes de taille à la naissance a été fixée à un centimètre. En fait, on a gardé la précision maximale.

La variation de taille depuis la naissance a été obtenue en retirant bien évidemment la valeur initiale (à la naissance) de celle mesurée à la fin de l'observation. Ce calcul fournit une variation qui correspondrait à un âge par différence de millésimes dans le cas classique. À titre d'exemple, les individus mesurant respectivement 50 et 53 centimètres à la naissance et à la fin de l'observation ont enregistré une variation de taille de 3 centimètres en différence de tailles. Comme le montre la figure 3.9, cette variation recouvre en fait des gains de taille allant de 2 à 4 centimètres exacts.

Figure 3.9. Variation de taille en « différence de millésimes »



Cette figure permet aussi de comprendre pourquoi nous avons décidé de considérer les tailles en centimètres révolus. En les considérant comme des tailles exactes, toutes les sorties d'observation d'individus de taille de naissance 50 centimètres et de taille de sortie d'observation égale à 53 centimètres se regrouperaient en un seul point de coordonnées 53 en abscisse et 3 en ordonnée. Les observations se regrouperaient à chaque fois sur des intersections de droites, le reste du graphique n'en comportant aucune.

3.3.2.2.2. Observations retenues pour l'élaboration des graphiques

Avant d'opérer le classement des données en fonction de la taille à la naissance et de la variation de taille depuis la naissance, il a été décidé de filtrer les données. Tout d'abord, les individus dont la taille de naissance était soit strictement inférieure à 40 centimètres, soit strictement supérieure à 55 centimètres ont été retirés. La cohorte de 39 centimètres ne compte que 124 individus et pour les valeurs inférieures, les chiffres décroissent très vite (cf. annexe 3.3). La cohorte de 56 centimètres n'en compte que 194 et à partir de là, les effectifs décroissent très rapidement.

Pour la taille finale, la limite inférieure a été fixée à 39 centimètres de façon à ne pas avoir de taille finale inférieure à la taille initiale et la limite supérieure à 90 centimètres. Cette limite a été choisie notamment parce que les cas deviennent très rares pour ces valeurs et surtout parce

¹⁷ On aurait aussi pu considérer que les données avaient été arrondies, mais cette option aurait compliqué inutilement la suite des opérations, dont l'établissement des diagrammes de Lexis.

que la valeur 99 est une des façons (avec le 0) d'indiquer une donnée manquante ; certaines valeurs dans la dizaine des 90 pourraient être en fait des données manquantes mal encodées. Finalement, pour la variation de taille, seules les valeurs strictement supérieures à 0 ont été retenues pour éviter le problème du premier triangle¹⁸. L'application croisée de ces trois critères laisse 82 523 observations qui se répartissent comme indiqué dans le tableau 3.3.a.

La cohorte de taille à la naissance la moins nombreuse compte 180 individus (40 centimètres), et la plus nombreuse, 16 906 (50 centimètres). Les effectifs à chaque variation de taille ont été divisés par l'effectif des cohortes concernées pour obtenir les événements réduits. Ce sont ces derniers, exprimés en « pour dix mille » qui serviront pour établir les diagrammes de Lexis (*cf.* tableau 3.3.b).

Tableau 3.3. Observations classées selon la taille à la naissance et la variation de taille

a. Les événements observés

Gain de taille	Taille à la naissance																Total
	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	
1					1		9	7	13	14	10	10	9	4	2		79
2			1		3	1	6	5	23	17	29	19	8	5	2		119
3				1	2	2	11	13	19	29	28	24	13	4			146
4		1		1		3	5	18	18	22	31	25	6	8	8	1	147
5			1	1	3	4	7	19	22	28	32	31	13	6	3	1	171
6		1		1	2	1	9	21	21	37	38	39	23	6	4		203
7		1		1	2	6	10	16	26	44	50	41	16	18	3	4	238
8			2	1	2	7	2	16	36	44	53	35	23	13	7	3	244
9			1		1	2	11	25	32	40	54	39	29	10	7	2	253
10			2	1	2	2	9	16	36	47	64	37	29	17	4	2	268
11				1	1	6	15	24	38	53	51	56	24	12	4	1	286
12	1	2		1	3	5	10	24	41	60	65	51	34	16	5	2	320
13			2	1	6	4	10	24	31	36	55	40	33	13	9		264
14	2		1	3	4	5	8	29	37	50	50	59	35	20	5	4	312
15	1		1	2	3	7	13	23	28	52	59	43	47	24	10	5	318
16			3		5	6	17	34	53	65	67	52	46	21	13	5	387
17	1		1	4	7	9	19	33	50	69	99	80	54	51	16	12	505
18	3		1	1	3	9	23	41	76	103	133	134	108	46	41	22	744
19		3	3	5	4	14	25	51	101	192	243	248	167	113	62	26	1 257
20	1		4	3	7	22	55	101	196	297	493	351	345	188	111	32	2 206
21	1		2	4	20	36	90	166	316	558	805	704	481	309	129	57	3 678
22	1	1	3	8	28	53	128	303	658	912	1 375	1 017	759	406	188	72	5 912
23	1	1	6	13	20	73	231	565	940	1 442	1 901	1 387	972	443	192	68	8 255
24	3	3	10	24	40	115	363	666	1 326	1 830	2 379	1 680	1 075	494	173	38	10 219
25	6	4	18	35	78	192	464	936	1 560	2 038	2 502	1 797	925	464	180	51	11 250
26	5	7	24	50	99	244	549	1 008	1 631	2 025	2 192	1 390	811	310	138	29	10 512
27	10	13	23	60	141	317	515	978	1 374	1 635	1 593	1 095	536	232	73	20	8 615
28	16	21	52	65	139	269	450	787	1 000	1 094	1 143	651	368	135	48	16	6 254
29	15	38	40	69	123	222	346	548	656	709	652	411	174	77	21	7	4 108
30	23	27	41	61	106	162	260	350	415	389	370	167	101	27	7	2	2 508
31	18	28	39	57	83	108	110	178	214	199	166	85	33	13	3	3	1 337
32	16	24	30	43	36	71	74	104	104	78	65	38	18	4	1	1	707
33	14	25	30	31	21	34	43	47	32	41	30	18	5		3		374
34	16	12	14	9	10	13	13	15	15	9	16	9	2		1	1	155
35	9	7	10	9	4	8	5	9	6	4	6	3		1			81
36	12	8	3	2	3	1	1	5	1	1	4	3	2	1			47
37	3	5	1			3	1	3		1	2		1				20
38	1	1		2	1	1		1	1	1		2					11
39			1					1	1	1	1						5
40	1							1		1							3
41			1														3
43		1	1					2									2
Total	180	234	372	570	1 013	2 037	3 917	7 213	11 147	14 267	16 906	11 871	7 325	3 511	1 473	487	82 523

¹⁸ Comme montré en annexe 3.5, il serait possible de travailler avec les triangles dans le cas des diagrammes en taille, mais, outre le fait que le travail de préparation des données est beaucoup plus lourd, la représentation des données n'est pas plus précise vu que, au départ, le calcul s'opère en différence de « millésimes ». Le peu d'intérêt de la démarche dans la représentation des données comparé à la lourdeur des opérations nous a amené à privilégier les diagrammes en différence de millésimes sans premier triangle pour le cas de la taille.

b. Les évènements réduits (en pour dix mille)

Gain de taille	Taille à la naissance															
	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55
1	0	0	0	0	10	0	23	10	12	10	6	8	12	11	14	0
2	0	0	27	0	30	5	15	7	21	12	17	16	11	14	14	0
3	0	0	0	18	20	10	28	18	17	20	17	20	18	11	0	0
4	0	43	0	18	0	15	13	25	16	15	18	21	8	23	54	21
5	0	0	27	18	30	20	18	26	20	20	19	26	18	17	20	21
6	0	43	0	18	20	5	23	29	19	26	22	33	31	17	27	0
7	0	43	0	18	20	29	26	22	23	31	30	35	22	51	20	82
8	0	0	54	18	20	34	5	22	32	31	31	29	31	37	48	62
9	0	0	27	0	10	10	28	35	29	28	32	33	40	28	48	41
10	0	0	54	18	20	10	23	22	32	33	38	31	40	48	27	41
11	0	0	0	18	10	29	38	33	34	37	30	47	33	34	27	21
12	56	85	0	18	30	25	26	33	37	42	38	43	46	46	34	41
13	0	0	54	18	59	20	26	33	28	25	33	34	45	37	61	0
14	111	0	27	53	39	25	20	40	33	35	30	50	48	57	34	82
15	56	0	27	35	30	34	33	32	25	36	35	36	64	68	68	103
16	0	0	81	0	49	29	43	47	48	46	40	44	63	60	88	103
17	56	0	27	70	69	44	49	46	45	48	59	67	74	145	109	246
18	167	0	27	18	30	44	59	57	68	72	79	113	147	131	278	452
19	0	128	81	88	39	69	64	71	91	135	144	209	228	322	421	534
20	56	0	108	53	69	108	140	140	176	208	292	296	471	535	754	657
21	56	0	54	70	197	177	230	230	283	391	476	593	657	880	876	1 170
22	56	43	81	140	276	260	327	420	590	639	813	857	1 036	1 156	1 276	1 478
23	56	43	161	228	197	358	590	783	843	1 011	1 124	1 168	1 327	1 262	1 303	1 396
24	167	128	269	421	395	565	927	923	1 190	1 283	1 407	1 415	1 468	1 407	1 174	780
25	333	171	484	614	770	943	1 185	1 298	1 399	1 428	1 480	1 514	1 263	1 322	1 222	1 047
26	278	299	645	877	977	1 198	1 402	1 397	1 463	1 419	1 297	1 171	1 107	883	937	595
27	556	556	618	1 053	1 392	1 556	1 315	1 356	1 233	1 146	942	922	732	661	496	411
28	889	897	1 398	1 140	1 372	1 321	1 149	1 091	897	767	676	548	502	385	326	329
29	833	1 624	1 075	1 211	1 214	1 090	883	760	588	497	386	346	238	219	143	144
30	1 278	1 154	1 102	1 070	1 046	795	664	485	372	273	219	141	138	77	48	41
31	1 000	1 197	1 048	1 000	819	530	281	247	192	139	98	72	45	37	20	62
32	889	1 026	806	754	355	349	189	144	93	55	38	32	25	11	7	21
33	778	1 068	806	544	207	167	110	65	29	29	18	15	7	0	20	0
34	889	513	376	158	99	64	33	21	13	6	9	8	3	0	7	21
35	500	299	269	158	39	39	13	12	5	3	4	3	0	3	0	0
36	667	342	81	35	30	5	3	7	1	1	2	3	3	3	0	0
37	167	214	27	0	0	15	3	4	0	1	1	0	1	0	0	0
38	56	43	0	35	10	5	0	1	1	1	0	2	0	0	0	0
39	0	0	27	0	0	0	0	1	1	1	1	0	0	0	0	0
40	56	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0
41	0	0	27	0	0	0	0	3	0	0	0	0	0	0	0	0
43	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total	10 000	10 000	10 000	10 000	10 000	10 000	10 000	10 000	10 000	10 000	10 000	10 000	10 000	10 000	10 000	10 000

(Source : Banque de Données médico-sociales de l'ONE ; cf. point 3.3.2.1.)

À titre d'exemple, parmi les 16 906 individus de la cohorte 50 centimètres, 1 593 ont enregistré une variation de taille à la sortie d'observation de 27 centimètres (ils avaient donc une taille en fin d'observation de 77 centimètres). Traduites en évènements réduits, ces observations donnent 942,27 pour dix mille (soit 1 593/1,6906).

3.3.2.2.3. Graphiques tous sexes et toutes durées

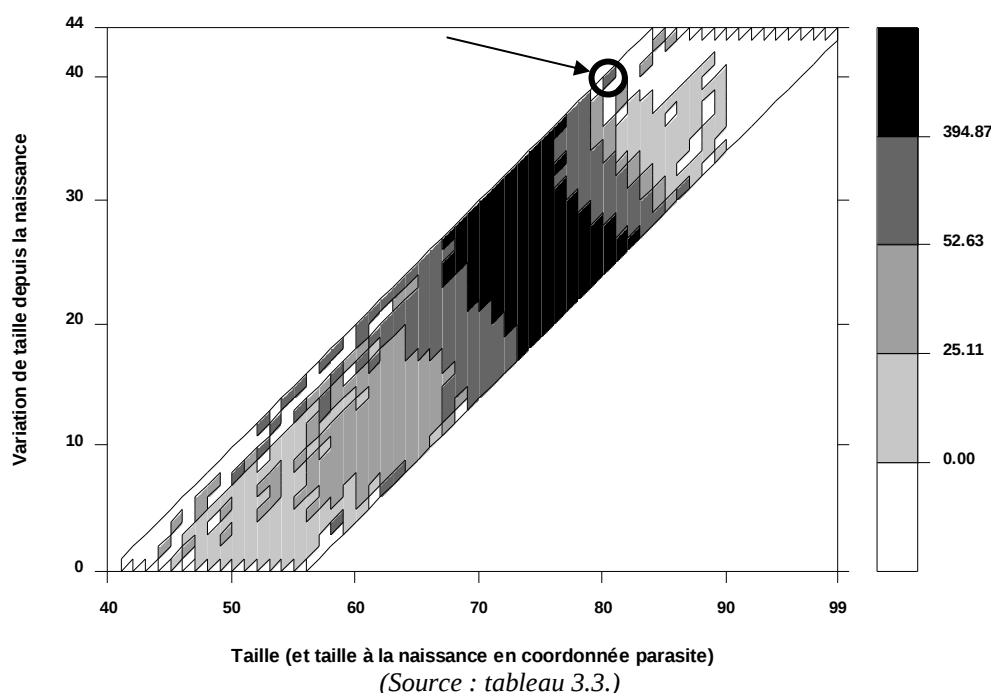
a. Choix des limites de classes

Un premier graphique a été réalisé selon la procédure suivante (cf. figure 3.10¹⁹) : les valeurs strictement nulles ont été retirée, soit 139 sur un total de 688 valeurs pour former une classe à part. Il restait donc 549 valeurs d'évènements réduits (ou parallélogrammes) à distribuer. Ils l'ont été sur la base de quartiles. Le premier quartile en reprend 137 et va jusqu'à 25,11 ; le deuxième, 136 jusqu'à 52,63 ; le troisième, 139 jusqu'à 394,87 et le dernier, 137. Comme le

¹⁹ Ce graphique – ainsi que tous les suivants – a été élaboré sur base du programme « Lexis », mis au point par K. Andreev. Ce programme est d'accès libre (renseignements auprès de K. Andreev, adresse : andreevk@post.queensu.ca ou auprès du Max Planck Institute for Demographic Research, adresse : <http://www.demogr.mpg.de/>). La figure 3.10 a été construite sur base d'un tableau croisé SPSS montrant la distribution des évènements en fonction de la cohorte de taille à la naissance et de la variation de taille à la sortie de l'observation ; ce tableau croisé a été importé en Excel. Les évènements réduits calculés au départ de ces données ont été arrangés adéquatement sur une feuille Excel qui a été ensuite sauvée au format TXT. Moyennant quelques petits traitements, ce fichier TXT a été importé dans le logiciel Lexis.

montrent ces chiffres, le fait d'avoir rejeté les valeurs nulles avant le calcul des quartiles nous a pratiquement donné des valeurs de quintiles sur les 688 valeurs initiales²⁰ !

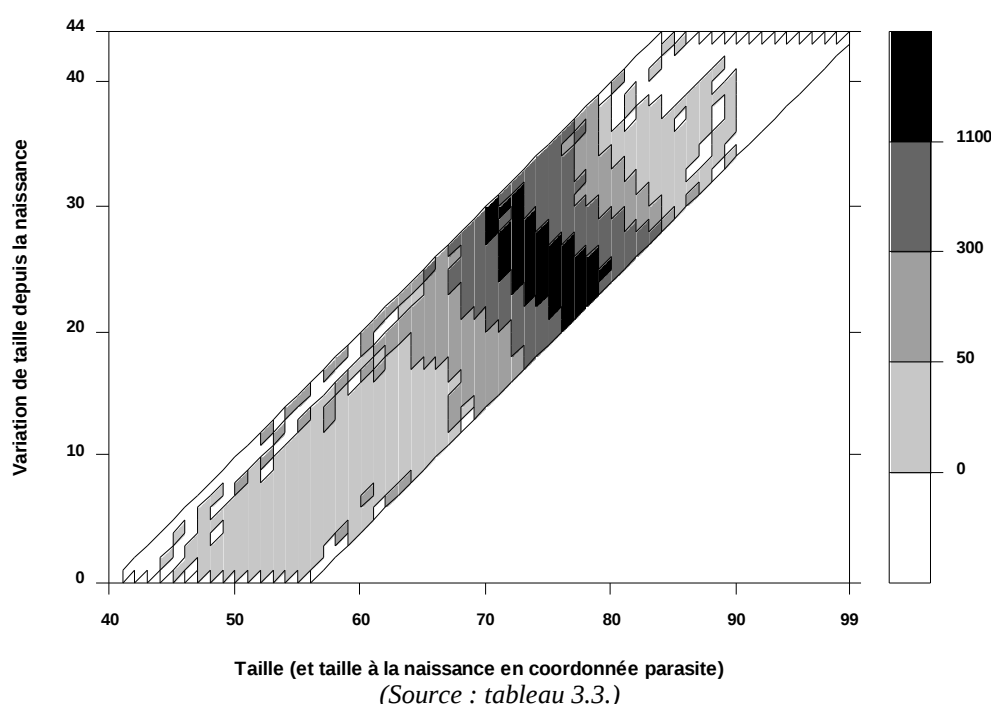
Figure 3.10. Taille à la fin de l'observation (en cm) – Tous sexes – Quartiles (événements réduits en pour dix mille)



Dans cette figure, le parallélogramme entouré d'un cercle et indiqué par une flèche représente les événements réduits de la cohorte de 40 centimètres à la naissance pour un gain de taille de 40 centimètres. Cette cohorte est celle qui présente la taille la plus faible à la naissance ; elle occupe donc le premier couloir oblique à la gauche de la figure dont la taille à la naissance se lit sur l'axe horizontal. Le gain de 40 centimètres se lit sur l'axe vertical. Selon le tableau 3.3.b, dans cette cohorte et pour le gain de taille considéré, les événements réduits s'élèvent à 56 pour dix mille ; autrement dit le parallélogramme en cause tombe dans la classe allant de 52,63 à 394,87 pour dix mille.

Cette figure montre clairement une tendance dans l'organisation des données ; toutefois, avant d'aller plus loin dans le commentaire, nous avons préféré adopter d'autres limites de classe, notamment parce que la classe supérieure (au-delà de 394,87) occupe une place « centrale » par trop importante. Après différents essais, nous avons opté pour les limites suivantes (cf. figure 3.11) : 50, 300 et 1 100 pour dix mille. La deuxième classe compte 272 valeurs ; la troisième, 124 ; la quatrième, 101 et la cinquième 51. Ces effectifs sont certes très inégaux ; toutefois ce choix montre bien, à notre avis, l'organisation des observations et même mieux qu'avec les quartiles/quintiles (cf. figure 3.11).

²⁰ On aurait donc pu préférer calculer directement des quintiles. Toutefois, dans le cas du poids (cf. point suivant), l'application des quintiles aurait été plus délicate car la classe des valeurs nulles est trop importante : à elle seule elle dépasse les 20 % des valeurs. Dans ce cas, la procédure suivie (rejet des valeurs nulles avant calcul des quartiles) n'en est que plus indispensable.

Figure 3.11. Taille à la fin de l'observation (en cm) – Tous sexes (évè. réduits en pour dix mille)

La systématique de la distribution des événements réduits dans les différentes cohortes nous semble mieux mise en évidence visuellement sur ce deuxième graphique, notamment à l'intérieur des valeurs les plus fortes, initialement regroupées en une seule classe. Nous avons privilégié ce critère visuel et adopté ces limites de classe pour les autres graphiques en taille²¹.

b. Des effets de cohorte et de variation de taille

La figure 3.11 montre que les données s'organisent sous l'influence principale de deux effets :

- un premier effet de variation de taille (qui se dénommerait « effet d'âge » dans la configuration classique du diagramme de Lexis) : dans chacune des cohortes, la densité des événements varie en fonction du gain de taille depuis la naissance ; par ailleurs, la dispersion des événements épouse des schémas proches dans les différentes cohortes ;
- un deuxième effet de cohorte de taille à la naissance (dont la correspondance dans la configuration classique est immédiate) : à mesure que la taille à la naissance augmente, les événements se localisent pour des gains de taille de plus en plus faibles²².

Les trois caractéristiques signalées (effet de variation de taille, même schéma dans les différentes cohortes et effet de cohorte) peuvent être illustrées par quelques données complémentaires (*cf.* tableau 3.4). La variation moyenne décroît de manière continue à mesure que la taille à la naissance augmente : elle passe de 30,106 à 22,125 centimètres de la cohorte de 40 à celle de 55 centimètres. Le mouvement est relativement régulier. Ces chiffres attestent l'effet de cohorte : plus la taille à la naissance augmente, moins le gain de taille est important. L'écart type est plutôt constant, passant d'un minimum de 4,049 pour la cohorte 55

²¹ Il est à noter que passer à d'autres limites de classe s'opère très rapidement ; il suffit de modifier l'échelle sur le graphique et le logiciel Lexis exécute toutes les modifications nécessaires. Par ailleurs, nous aurions pu décider de construire les diagrammes de Lexis en privilégiant d'autres couples de variables, mais, à notre avis, la version de Brasche est la plus efficace (*cf.* annexe 3.6).

²² Cet effet pourrait être un effet de « régression sur la moyenne » ; à ce sujet, *cf.* annexe 3.4).

centimètres à un maximum de 4,992 pour la cohorte de 42 centimètres. Au sein des cohortes, la dispersion autour de la moyenne est fort semblable.

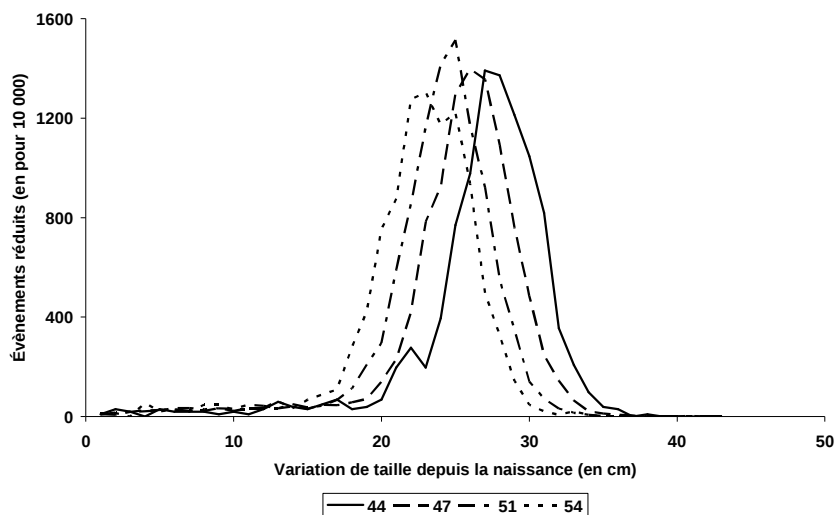
Tableau 3.4. Moyenne et écart-type par cohorte de taille

Taille à la naissance	Effectifs de la cohorte	Variation moyenne	Écart-type de la variation
40	180	30,106	4,664
41	234	29,983	4,607
42	372	28,513	4,992
43	570	27,909	4,476
44	1 013	26,914	4,663
45	2 037	26,605	4,207
46	3 917	25,677	4,459
47	7 213	25,330	4,421
48	11 147	24,857	4,280
49	14 267	24,470	4,228
50	16 906	24,110	4,135
51	11 871	23,727	4,262
52	7 325	23,353	4,099
53	3 511	22,928	4,125
54	1 473	22,544	4,146
55	487	22,125	4,049
Total	82 523	24,414	4,367

(Source : tableau 3.3.)

La figure 3.12 montre bien l'effet de la variation de taille sur la distribution (l'équivalent d'un effet d'âge) ; en définitive, il s'agit tout simplement de la dispersion des observations autour d'une valeur centrale. Pour les cohortes retenues, les courbes sont fort semblables ; elles se décalent progressivement vers la gauche à mesure que la taille à la naissance augmente. En définitive, ce graphique illustre bien à lui seul les trois caractéristiques signalées.

Figure 3.12. Courbe des événements réduits par variation de taille



(Source : tableau 3.3.)

Les figures 3.11 et 3.12 poursuivent le même objectif : représenter la distribution des événements réduits par cohorte de taille à la naissance et par variation de taille. Dans le cas de la figure 3.12, nous avons dû nous limiter à 4 cohortes ; en effet, à cause de l'enchevêtrement des courbes, si le graphique en avait contenu plus, il serait devenu illisible, ou du moins difficile à lire. Au contraire, la figure 3.11 reprend toutes les cohortes tout en mettant bien en évidence les caractéristiques essentielles des distributions d'événements. La différence vient

de ce que dans le diagramme de Lexis, un espace particulier est réservé à chacune des cohortes, ce qui n'est pas le cas dans la figure 3.11. Sans vouloir en rien nier l'intérêt de la figure 3.11, cette discussion tendrait à attribuer certaines qualités particulières à une représentation de données de ce type en recourant à un diagramme de Lexis.

Différents éléments pourraient interférer avec la distribution du phénomène, dont :

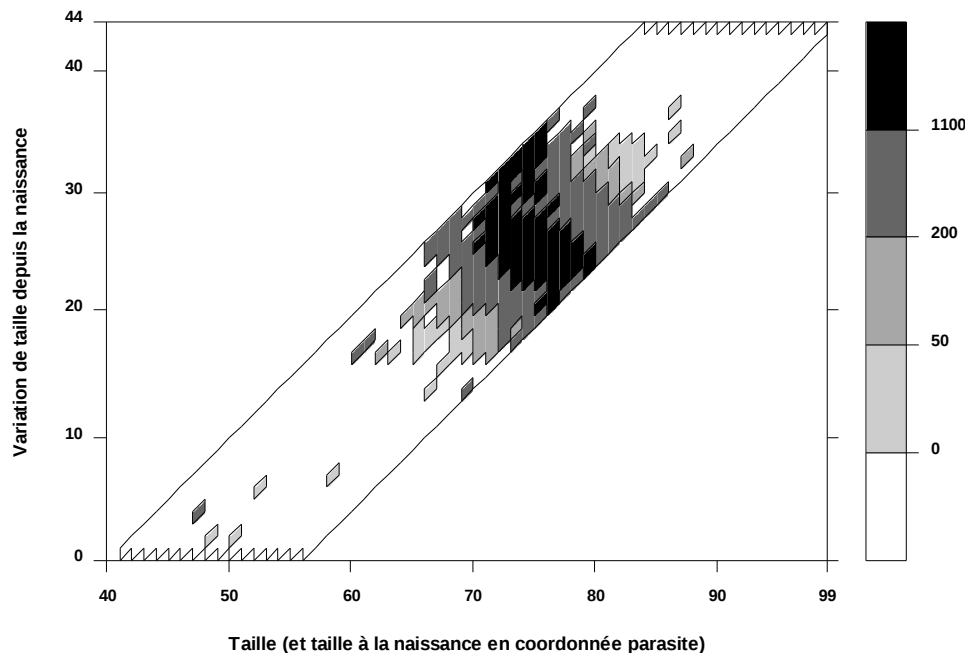
- la sortie d'observation peut se faire à des âges différents ;
- les données mélangent les deux sexes ; or, la taille varie en fonction du sexe ;
- les données mélangent des naissances survenues après des durées de gestation différentes ; il s'agit là aussi d'une cause possible de différenciation du comportement.

Les points qui vont suivre serviront à vérifier si le constat établi sur base de la figure 3.11 résiste à ces remarques.

3.3.2.2.4. Taille à un an et âge moyen en fin d'observation par cohorte

Dans un premier temps, on a sélectionné les seules observations pour lesquelles la dernière consultation a eu lieu approximativement au moment du premier anniversaire. Dans les données initiales, pour beaucoup d'observations, la date de la dernière consultation n'a pas été notée. En définitive, ce renseignement n'est disponible que pour les naissances des années 1999 et 2000. Par ailleurs, pour obtenir un nombre suffisant d'observations, il a été décidé d'assimiler au premier anniversaire toutes les durées d'observation comprises entre 51 et 52 semaines, soit 5 369 cas.

Figure 3.13. Taille à un an (en cm) – Tous sexes (événements réduits en pour dix mille)



(Source : Banque de Données médico-sociales de l'ONE ; cf. point 3.3.2.1.)

Comme le montre la figure 3.13, l'effet de cohorte reste bien visible : à mesure que la taille de naissance augmente, la gain de taille dans la première année de vie diminue. L'effet de variation de taille est moins net²³, mais toujours bien présent. Le comportement des cohortes semble moins homogène, mais il ne faut pas oublier ici les petits nombres en jeu (minimum : cohorte de 40 centimètres, 14 cas ; maximum : cohorte de 50 centimètres, 1 085 cas).

La figure 3.13 se singularise par rapport à la figure 3.11 par le fait que la dispersion des observations est moins importante : suite au contrôle de l'âge, elles se regroupent en grande majorité dans une zone centrale précise. Quelques rares cas se situent à des valeurs très faibles de variation de taille (8 observations pour une variation inférieure à 15 centimètres). Il s'agit sans doute d'erreurs d'encodage pour la plupart.

Enfin, le comportement erratique des cohortes extrêmes est à mettre en relation avec leurs (très) faibles effectifs, avec respectivement 14, 18, 27, 46 et 29 cas pour les cohortes de 40, 41, 42, 43 et 55 centimètres !

Des graphiques semblables à la figure 3.13 ont été établis pour des durées d'observation d'une part de 47 et 48 semaines, et, d'autre part, de 49 et 50 semaines. Ces deux figures – non proposées dans le texte – présentent les mêmes caractéristiques que la figure 3.13.

Il faut toutefois signaler un décalage logique vers des gains moyens de plus en plus élevés à mesure que la durée d'observation augmente (avec *grosso modo* une augmentation de 0,25 centimètre par semaine d'observation en plus²⁴ ; cf. tableau 3.5). Bref, le constat émis sur base de la figure 3.11 n'est pas contredit quand la durée d'observation est contrôlée.

Tableau 3.5. Durée d'observation et gain de taille

Durée d'observation	Effectifs sous observation	Gain de taille moyen
47 et 48 semaines	3 806	24,55 centimètres
49 et 50 semaines	5 006	25,00 centimètres
51 et 52 semaines	5 369	25,51 centimètres

(Source : Banque de Données médico-sociales de l'ONE ; cf. point 3.3.2.1.)

²³ Un autre choix de limites de classe pourrait le rendre plus apparent ; cet autre choix pour les limites pourrait aussi se justifier par une moins grande dispersion des événements suite au contrôle de l'âge. Toutefois, nous avons préféré conserver les mêmes limites pour l'ensemble des graphiques pour en assurer la comparabilité.

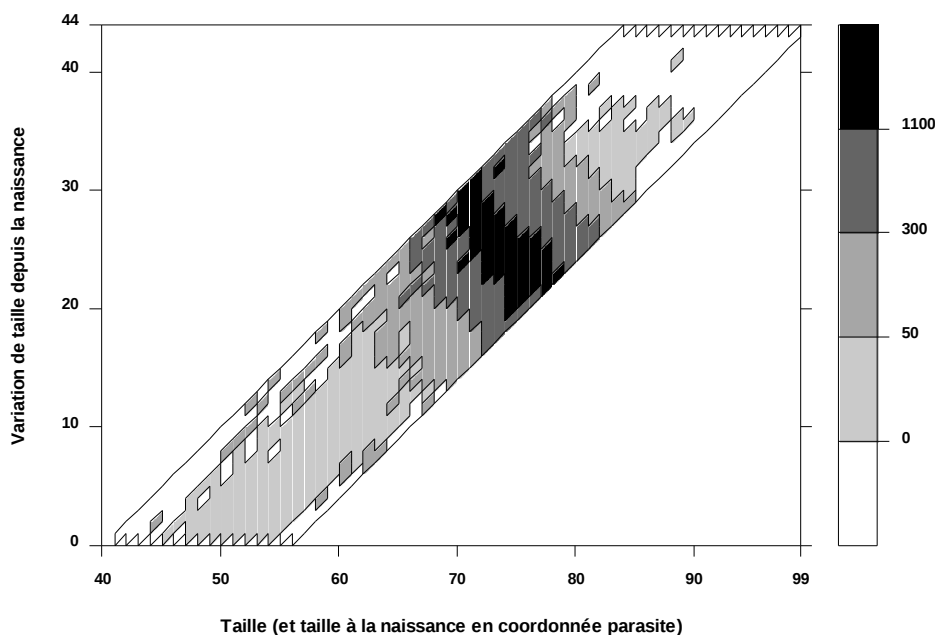
²⁴ Ce rythme de croissance est conforme aux courbes standard proposées dans la littérature. À ce sujet, cf. notamment R. Kuczmarski, C. Ogden, S. Guo *et al.* (2002) ; en p. 11, ces auteurs comparent notamment un standard ancien (The 1977 NCHS Growth Charts) avec le nouveau qu'ils proposent (The 2000 CDC Growth Curves). Un graphique comparant ces deux standards est proposé en p. 115 ; aucune différence notable ne les oppose. On pourra aussi consulter à ce sujet les courbes utilisées en Belgique francophone dans D. Lecleir (2003), pp. 72-73 ; il s'agit de courbes de croissance établies sur base d'un suivi longitudinal d'enfants durant la période 1955-1972.

3.3.2.2.5. Taille par sexe

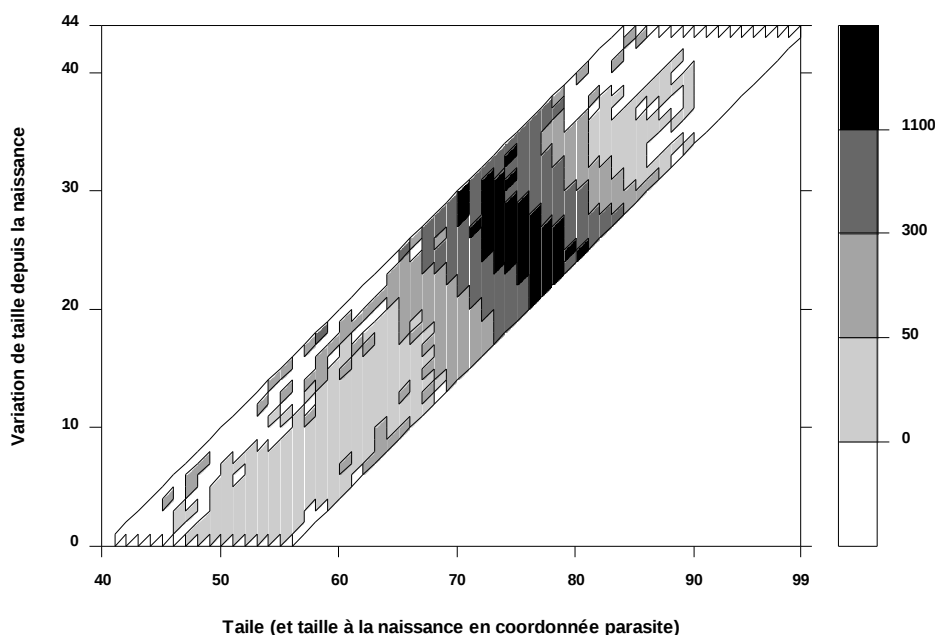
Le sexe est un facteur de différenciation du comportement en matière de taille. Ainsi, à la naissance, la taille moyenne parmi les 40 688 filles était de 48,96 centimètres contre 49,68 pour les 41 698 garçons²⁵ ; pour la fin de l'observation, les mêmes données étaient respectivement de 72,96 et 74,50. La figure 3.14 montre clairement que les deux sexes suivent en gros le même schéma de répartition des observations que le total général.

Figure 3.14. Taille à la fin de l'observation (en cm) – Sexes séparés (événements réduits en pour dix mille)

a. Sexe féminin



b. Sexe masculin



(Source : Banque de Données médico-sociales de l'ONE ; cf. point 3.3.2.1.)

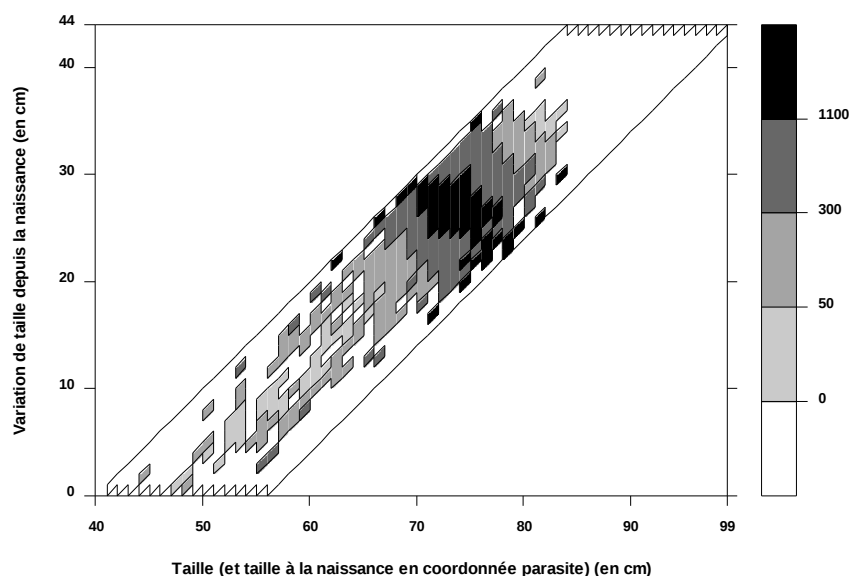
²⁵ Pour 137 naissances, le sexe est inconnu.

3.3.2.2.6. Taille par durée de gestation

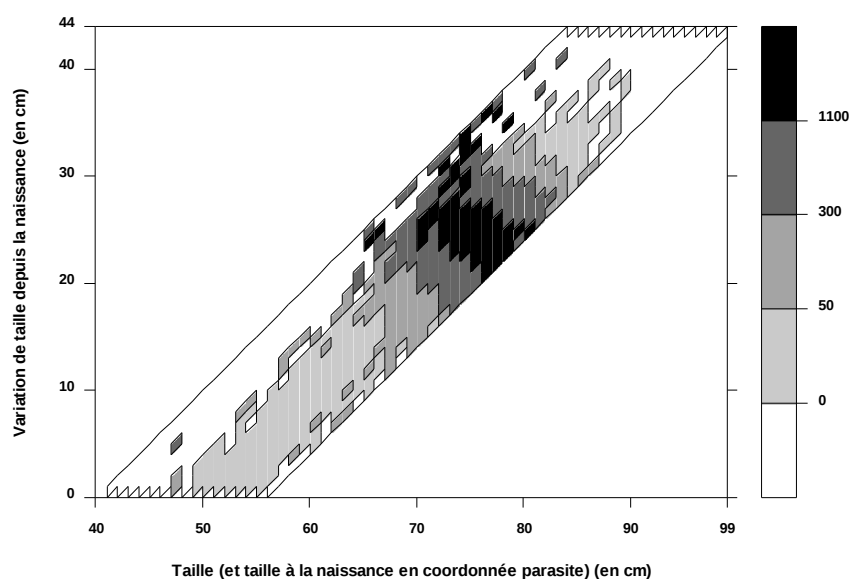
La figure 3.15 propose la distribution des évènements réduits pour les durées de gestation de 36 (2 763 cas) et de 40 semaines (26 022 cas). Les mêmes tendances générales se dessinent. Ainsi, à mesure que la taille à la naissance augmente, la variation de taille en fin d'observation diminue. Il est à noter que les cohortes à comportement erratique se constatent pour les valeurs élevées de taille à la naissance dans le cas des courtes durées de gestation (avec en plus la première cohorte) et de valeurs faibles pour les longues durées. Cela s'explique par le très faible nombre de cas dans ces cohortes : à 36 semaines, respectivement 5, 15, 6 et 2 observations pour les cohortes 40, 53, 54 et 55 centimètres et à 40 semaines, 19, 6 et 13, pour les cohortes 40, 41 et 42 centimètres.

Figure 3.15. Taille à la fin de l'observation (en cm) par durée de gestation (évènements réduits en pour dix mille)

a. 36 semaines de durée de gestation



b. 40 semaines de durée de gestation



(Source : Banque de Données médico-sociales de l'ONE ; cf. point 3.3.2.1.)

En résumé, cette analyse sur la variation de taille indique que les cohortes les plus grandes à la naissance grandissent moins dans la première année de leur vie que les cohortes de plus petite taille²⁶. Avant de revenir sur ce sujet d'un point de vue plus théorique, il nous a paru utile de montrer la distribution des événements en prenant, cette fois, les variables en poids comme cadre de référence.

3.3.2.3. Diagrammes en poids

3.3.2.3.1. Choix de l'évènement et calcul des coordonnées

La BDMS comprend également le poids à la naissance et à la sortie d'observation. En conséquence, il sera également possible d'élaborer des diagrammes de Lexis pour montrer la distribution des événements réduits du phénomène constitué par la sortie d'observation en fonction des trois variables de poids, poids à la naissance et variation de poids depuis la naissance. C'est donc dans un cadre de référence en poids, et plus en taille comme au point précédent, que l'analyse se déroulera ici.

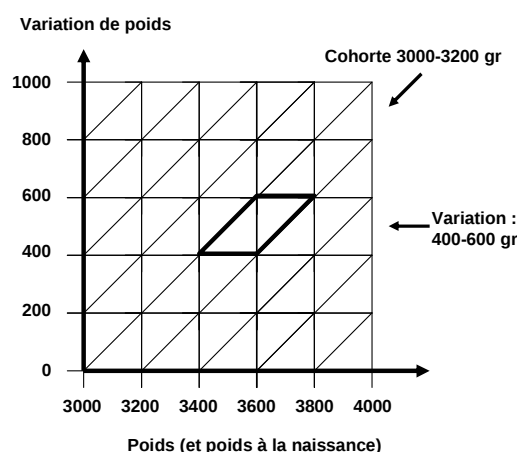
Les poids à la naissance et en fin d'observation ont été exprimés en multiples de 5 grammes, à de rares exceptions près, mais ils présentent une forte attraction pour les multiples de 10, encore mieux marquée pour ceux de 50 ou 100 (*cf.* annexe 3.7). Malgré cela, pour les calculs, nous avons considéré que les données étaient individuellement exactes et exprimées en grammes exacts. La variation de poids depuis la naissance a été obtenue bien évidemment en retirant la valeur à la naissance de celle à la fin de l'observation.

Pour éviter un trop grand nombre de cohortes, l'amplitude des cohortes a été fixée à 200 grammes pour le poids. En optant pour des cohortes de 200 grammes, la cohorte la plus nombreuse (entre 3 200 et 3 400 grammes) compte un effectif de 19 927, soit le même ordre de grandeur que dans le cas de la taille (pour rappel, la cohorte de 50 centimètres compte 22 496 observations).

Le poids final et la variation de poids ont aussi été regroupés par classes de 200 grammes. Même si ce choix ne s'imposait pas, nous avons opté pour une représentation des données de poids en fonction du poids de naissance et de la variation de poids depuis la naissance. De la sorte, les données se distribuent dans des parallélogrammes reposant sur une base²⁷. Ainsi, sur la figure 3.16, le parallélogramme correspondant à une variation de poids comprise entre 400 et 600 grammes pour des individus de la cohorte 3000 à 3200 grammes à la naissance a été mis en évidence.

²⁶ Il est à noter que nous avons aussi produit des graphiques en contrôlant l'âge des mères (par groupes quinquennaux), la parité, la situation par rapport au tabagisme. Le constat reste toujours le même, à quelques détails près. Il n'a pas été jugé utile de tous les reproduire, d'autant que nous reviendrons sur cette question dans le chapitre 4. Nous aurions voulu aussi croiser toutes ces caractéristiques, mais le contrôle d'une seule à la fois nous a déjà confronté régulièrement au piège des petits nombres...

²⁷ Dans le cas de la taille, le choix a été différent, mais surtout forcé vu la nature des données. Nous avons préféré travailler ici entre valeurs exactes de la variation de poids. De toutes façons, le choix de parallélogrammes sur pointe ou sur base ne change pratiquement rien à l'impression visuelle qui se dégage d'une distribution (*cf.* annexe 3.8).

Figure 3.16. Diagramme en poids avec parallélogrammes sur base

3.3.2.3.2. Observations retenues pour les différents graphiques

Pour la variation du poids, il a fallu inclure des valeurs négatives pour tenir compte de la perte de poids suivant la naissance. Cette perte peut aller jusqu'à 10 % du poids initial avec une moyenne de l'ordre de 6 %²⁸). La cohorte au poids de naissance le plus élevé retenue va jusqu'à 4 600 grammes. Le choix de la limite de – 800 pour la variation représente donc une marge de sécurité. L'observation des effectifs aurait pu nous conduire à une autre limite (– 600 grammes) et à considérer qu'au-delà il s'agissait d'erreur d'encodage. Par (excès de) sécurité, nous avons conservé la limite de – 800 grammes. Finalement, les données ont été filtrées selon les critères suivants :

- poids de naissances compris entre 2 000 et 4 600 grammes ;
- poids final compris entre 1 600 et 18 000 grammes ;
- variation de poids comprise entre – 800 et 16 000 grammes.

Ces différents choix ont été posés afin d'éviter de trop petits nombres (cf. tableau 3.6). Le croisement des différents critères laisse 89 392 observations, réparties en 13 cohortes. Après ce filtrage, la cohorte la plus nombreuse (3 200 à 3 400 grammes) compte 15 667 observations et la moins nombreuse (4 400 à 4 800), 522. Les effectifs à chaque variation de poids ont été divisés par celui de la cohorte correspondante pour obtenir les événements réduits. À titre d'exemple, parmi les 3 017 individus de la cohorte 4 000 à 4 200 grammes, un est sorti d'observation pour une variation comprise entre – 800 et – 600 grammes, ce qui donne 3,31 pour dix mille en termes d'événements réduits (soit 1/0,3017).

3.3.2.3.3. Graphique tous sexes et toutes durées

a. Choix des limites de classe

Un premier graphique a été établi sur la base des quartiles après retrait des valeurs nulles (avec les limites suivantes : 0 ; 17,59 ; 41,81 et 284,12). Toutefois, afin de mieux mettre en évidence ce qui se passe au centre de la distribution, nous avons opté, après différents essais, pour d'autres limites (0 ; 25 ; 200 et 525). Celles-ci ont été adoptées définitivement pour tous les graphiques en poids. Ce choix de limites est donc arbitraire ; toutefois, un autre choix ne devrait pas remettre en cause les analyses qui seront faites.

²⁸ Cf. D. Secker (1999), p. 3.

Tableau 3.6. Observations classées selon le poids à la naissance et la variation de poids

(VP = variation de poids entre la naissance et la sortie d'observation ; - 800 signifie de - 800 -< -600.

PI = poids à la naissance ; 2 000 = 2 000 -< 2 200.)

VP\PI	2000	2200	2400	2600	2800	3000	3200	3400	3600	3800	4000	4200	4400	Total
-800											1			1
-600														0
-400					1	1	3	3	5		1			14
-200			1	2	2	5	7	6	9	7	2	1		42
0		3	6	16	10	31	24	18	17	9	7	2		143
200	1		4	10	19	20	11	23	7	7	4	2		108
400	2		6	16	20	23	21	26	16	9	2	3		144
600	2	1	5	11	16	25	24	21	19	5	3	3		135
800	1	3	8	18	19	26	33	22	19	5	4	2		160
1000	1	3	7	10	23	23	23	31	21	19	2	2	1	166
1200	1	4	7	14	18	26	31	34	27	12	5	4	1	184
1400	1	4	10	23	34	38	36	41	20	18	5	3	1	234
1600	1	4	6	19	28	37	39	41	23	17	6	5	2	228
1800	4	4	5	13	21	42	36	41	23	13	7	5	1	215
2000	5	5	11	25	34	59	48	35	30	18	12	2	1	285
2200	2	3	10	17	29	40	48	45	31	11	12	3	5	256
2400	2	5	20	22	32	39	53	50	34	15	8	3	2	285
2600	3	8	14	21	29	38	46	38	30	19	7	8	1	262
2800	3		14	20	34	53	47	40	38	21	8	8	2	288
3000	5	3	9	41	34	36	65	35	34	25	13	2	2	304
3200	2	8	15	23	53	48	48	53	28	26	13	4		321
3400	5	5	17	27	30	60	57	47	36	36	13	5	2	340
3600	4	12	14	30	41	69	75	65	51	30	19	7	4	421
3800	5	16	23	30	58	89	87	79	56	39	23	12	5	522
4000	6	11	35	38	74	108	122	94	79	52	26	17	8	670
4200	10	19	28	76	118	143	147	141	129	73	55	16	11	966
4400	15	31	68	98	185	204	241	205	147	107	58	27	17	1 403
4600	15	38	68	133	204	309	323	323	210	133	97	28	13	1 894
4800	26	48	109	209	300	461	462	427	300	163	101	37	25	2 668
5000	30	55	138	246	439	577	622	550	343	253	120	75	18	3 466
5200	40	85	177	355	481	696	779	572	512	268	165	69	32	4 231
5400	45	85	202	368	631	850	909	838	583	316	172	80	32	5 111
5600	74	129	231	473	721	931	1 029	845	598	407	159	86	32	5 715
5800	73	121	270	534	770	1 060	1 066	944	657	372	219	78	30	6 194
6000	72	146	274	553	876	1 086	1 136	1 003	645	401	220	87	33	6 532
6200	83	134	296	521	868	1 044	1 167	940	684	390	202	107	26	6 462
6400	77	144	322	516	735	1 040	1 083	1 007	675	391	181	73	39	6 283
6600	72	158	264	453	707	943	1 074	869	632	376	190	70	27	5 835
6800	86	142	250	394	569	905	917	757	517	295	166	67	28	5 093
7000	51	106	197	366	587	783	755	657	455	305	119	60	22	4 463
7200	66	104	177	344	491	604	656	535	404	240	113	34	14	3 782
7400	55	87	160	239	374	469	565	500	328	196	108	34	19	3 134
7600	44	69	114	199	288	409	444	361	299	163	90	38	15	2 533
7800	39	61	100	183	240	323	340	301	234	130	84	35	15	2 085
8000	25	41	75	139	202	238	248	220	157	119	55	15	11	1 545
8200	21	28	60	103	131	187	201	177	159	62	34	24	7	1 194
8400	14	26	41	62	87	100	153	160	93	53	25	10	9	833
8600	13	20	26	41	70	92	112	102	69	53	20	19	4	641
8800	10	9	21	35	67	71	76	64	53	39	22	8	2	477
9000	4	9	12	32	57	56	44	56	29	15	11	7	2	334
9200	14	5	7	16	19	36	45	24	22	27	7	4		226
9400	2	6	7	13	17	19	28	21	18	12	6	3		152
9600		3	7	13	8	14	16	22	15	5	5	3		111
9800	1	4	7	2	10	16	13	11	9	4	2	2		81
10000		1	3	5	6	11	10	10	7	6	1		1	61
10200	1		1	1	7	7	7	4	4	3	2			37
10400	1	2	2	1	5	3	2	4	3	4	1			28
10600		1	2	1	2	4	5	5	3	2	1			26
10800	2		1	2	1	2	3	5	3		2			21
11000		1		1	2	3	3	1	4	1	1			17
11200		1		1	1	3		1	3					10
11400									1	1				2
11600														0
11800				1	1	1		1		1				5
12000					1		1	2						4
12200									2					2
12400					2									2
12600														0
12800														0
13000								1		1				2
13200								1						1
13400														0
13600			1				1							2
Total	1 137	2 021	3 965	7 175	10 939	14 636	15 667	13 555	9 659	5 800	3 017	1 299	522	89 392

(Source : Banque de Données médico-sociales de l'ONE ; cf. point 3.3.2.1.)

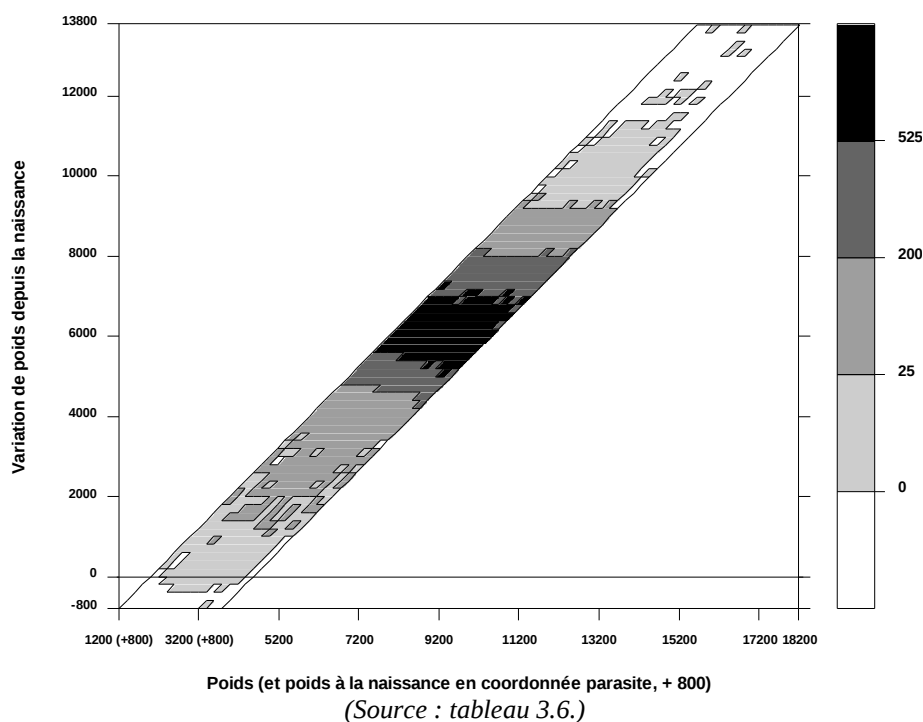
b. Des effets de variation, mais point ou si peu de cohorte

Avant de commenter la figure 3.17, il convient d'attirer l'attention sur l'axe des abscisses et spécialement sur ses étiquettes. En effet, vu la présence de valeurs négatives de variation du poids depuis la naissance, l'axe des ordonnées doit comporter des valeurs négatives (jusque - 800 grammes). Toutefois, dans le programme « Lexis », nous n'avons pas trouvé la possibilité de localiser l'échelle de l'axe horizontal ailleurs qu'à hauteur de l'extrémité inférieure de l'axe des ordonnées, soit à - 800 grammes.

En conséquence, la valeur du poids à la naissance correspond à la valeur du poids (en tant que variable absolue) augmentée de 800 grammes. Par exemple, à un poids (absolu) de 1 200 grammes correspond un poids de naissance de 2 000 grammes. L'indication « + 800 » n'a pas été reproduite au-delà de 3 200 grammes de poids (absolu) étant donné qu'il n'y avait pas de cohorte correspondant à l'étiquette suivante sur l'axe, soit 5 200 (la dernière cohorte prise en compte dans le graphique est celle qui va de 4 400 à 4 600 grammes). Tous les graphiques en poids seront construits selon ce principe.

La figure 3.17 montre l'équivalent d'un effet d'âge dans le diagramme classique : dans une cohorte, la densité des événements réduits varie en fonction de la variation de poids. De plus, toutes les générations suivent un schéma de distribution très semblable. Au contraire des diagrammes en taille, l'effet de cohorte est presque totalement absent : quelle que soit la cohorte, c'est toujours pour approximativement la même gamme de variation de poids que les fortes densités s'enregistrent ; toutefois, une légère tendance à la diminution du gain de poids se dessine à mesure que le poids de naissance augmente.

Figure 3.17. Poids en fin de l'observation (en grammes) – Tous sexes (événements réduits en pour dix mille)



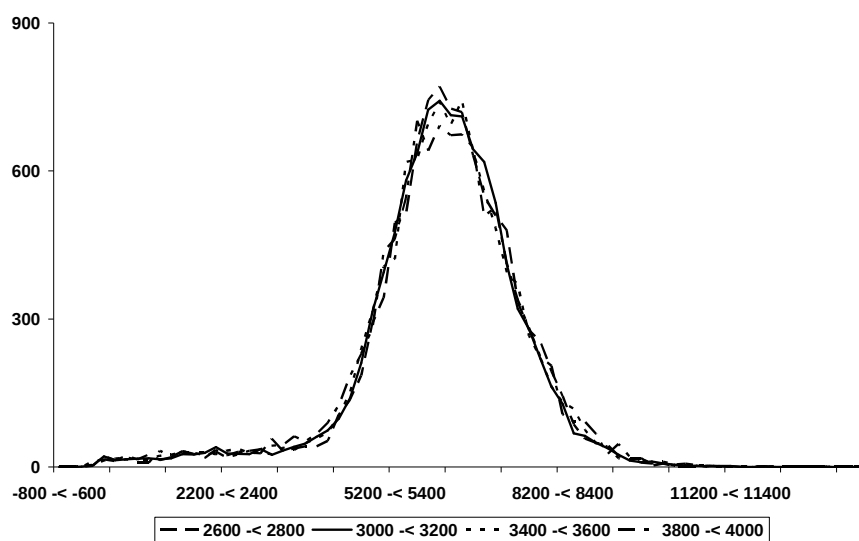
Le tableau 3.7 confirme bien ces constats : d'une part, la variation moyenne décline très légèrement, mais régulièrement, quand le poids de naissance augmente ; par ailleurs, l'écart type ne se modifie pratiquement pas d'une cohorte à l'autre.

Tableau 3.7. Variation moyenne et écart-type par cohorte de poids

Poids à la naissance	Effectifs de la cohorte	Variation moyenne	Écart-type de la variation
2 000-<2 200	1 137	6 323,307	1 375,756
2 200-<2 400	2 021	6 241,663	1 335,467
2 400-<2 600	3 965	6 114,497	1 388,378
2 600-<2 800	7 175	6 059,005	1 409,396
2 800-<3 000	10 939	6 052,195	1 383,380
3 000-<3 200	14 636	6 027,214	1 391,150
3 200-<3 400	15 667	6 041,824	1 380,181
3 400-<3 600	13 555	6 025,133	1 426,882
3 600-<3 800	9 659	6 026,961	1 443,893
3 800-<4 000	5 800	6 019,390	1 444,855
4 000-<4 200	3 017	5 988,320	1 421,863
4 200-<4 400	1 299	5 943,172	1 476,932
4 400-<4 600	522	5 940,230	1 375,223
Total	89 392	6 043,980	1 406,060

(Source : tableau 3.6.)

La figure 3.18 montre aussi le peu de différence entre les distributions dans les différentes cohortes retenues. La répartition par variation de poids suit le même schéma dans toutes les cohortes ; aucun décalage notable ne permet de bien distinguer les courbes. La situation est donc bien différente de ce que nous avons constaté pour le diagramme en taille, où un effet de cohorte était nettement visible.

Figure 3.18. Courbes des événements réduits par variation de poids

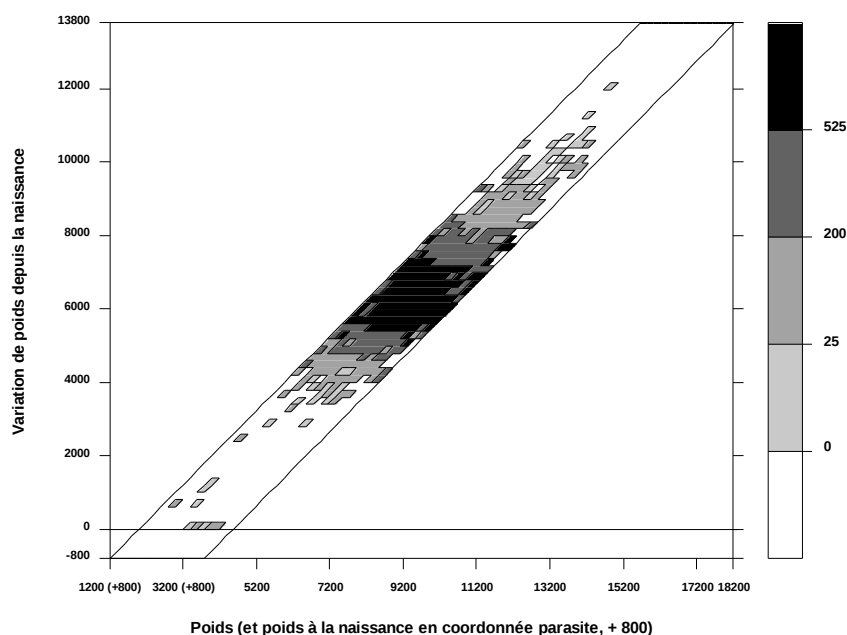
(Source : tableau 3.6.)

Comme signalé pour les diagrammes en taille, les figures 3.17 et 3.18 poursuivent le même objectif : représenter la densité des événements réduits par cohortes et par variation de poids. Il est à noter ici que la figure 3.18 ne laisse pas deviner le petit effet de cohorte alors que le diagramme de Lexis le laisse entrevoir. Outre une meilleure individualisation des cohortes (pas de croisement des courbes), ce dernier permettrait donc de mieux voir les choses. Il s'agit incontestablement d'une qualité en faveur de ce type de graphique.

Comme le montre la figure 3.19, cette description reste valable en contrôlant l'âge à la fin de l'observation (sélection des âges de 51 et 52 semaines, soit 5 798 cas). Le même constat s'impose en cas de contrôle de l'âge des mères, de la durée de gestation, de la parité ou de la situation de la mère par rapport au tabagisme (les figures y correspondant n'ont pas été

reproduites ici). Bref, le constat concernant l'effet de variation non accompagné d'un effet de cohorte marqué résiste à des analyses plus précises²⁹.

Figure 3.19. Poids à un an (en grammes) – Tous sexes (évè. réduits en pour dix mille)



(Source : Banque de Données médico-sociales de l'ONE ; cf. point 3.3.2.1.)

3.3.2.4. Diagrammes en poids ou en taille et effets : bilan

Les deux points précédents ont établi que la répartition des événements réduits subissait l'équivalent d'un effet d'âge aussi bien dans le cas de la taille que dans celui du poids ; par contre, si un effet de cohorte apparaît clairement sur les diagrammes en taille, ceux en poids n'en montrent – pratiquement – pas³⁰. Comment rendre compte de ces caractéristiques ?

La croissance en taille ou en poids des nourrissons dépend d'un grand nombre de facteurs aléatoires de natures diverses concernant aussi bien des caractéristiques individuelles (une tendance, éventuellement en partie génétique, à être plus ou moins gros ou grand) que d'autres « environnementales » liées aux conditions de vie (comme le régime alimentaire, par exemple³¹). Dans ces circonstances, il n'est pas étonnant que, au sein des cohortes, le gain de poids ou de taille se distribue avec une certaine variabilité autour d'une valeur centrale. L'effet « d'âge », présent aussi bien dans les diagrammes en taille qu'en poids, traduit simplement cette variabilité. Il est à noter que cet effet se retrouve dans toutes les cohortes ; il est donc indépendant de la valeur de la variable de référence à la naissance.

L'effet de cohorte trouvé dans le cas de la taille porte donc uniquement sur une différence de tendance centrale : à mesure que la taille à la naissance augmente, la variation centrale de taille s'amenuise. Du côté du poids, rien de semblable n'apparaît. Nous reviendrons plus loin sur ce constat.

²⁹ Les graphiques en poids et en taille ne reprennent pas exactement les mêmes observations suite aux filtrages. Des graphiques ont été exécutés sur les données doublement filtrées. Les conclusions demeurent.

³⁰ Ce constat se retrouve dans certaines recherches sur le sujet. À ce sujet, cf. notamment A. Lartay *et al.* (2000), p. 45 où il est montré que la vitesse de croissance en taille entre 1 et 6 mois est négativement corrélée avec la taille à 1 mois alors que du côté du poids, on ne retrouve pas cette corrélation.

³¹ À ce sujet, cf. notamment D. Secker (1999), p. 4.

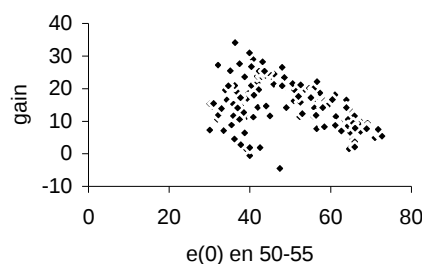
3.3.3. Des diagrammes de Lexis qui s'ignorent

Dans une partie précédente de ce texte (cf. point 2.3), nous avons montré que, si le diagramme de Lexis est bien un diagramme cartésien, il n'en présente pas moins des particularités qui débouchent sur la possibilité de cohortiser les phénomènes pour les mesurer. Dans la littérature, on peut rencontrer des diagrammes cartésiens qui en fait mériteraient d'être qualifiés « de Lexis » sans que la cohortisation ne soit mise à profit, ce qui pourrait, le cas échéant, améliorer la qualité de certaines mesures (cf. le point 5.5).

Dans ce point, nous allons partir de figures extraites de la publication suivante : Caselli G., Meslé F. et Vallin J., *Les entorses au schéma de la transition épidémiologique*. Communication présentée à la séance 18 : « Nouvelles menaces sanitaires », du Congrès international de la population. Salvador. Brésil. Août 2002, Rome/Paris, 44 p. En fait, les figures reprises ci-après ont été légèrement modifiées par rapport aux figures originales sans que cela ne remette en cause le raisonnement qui sera suivi.

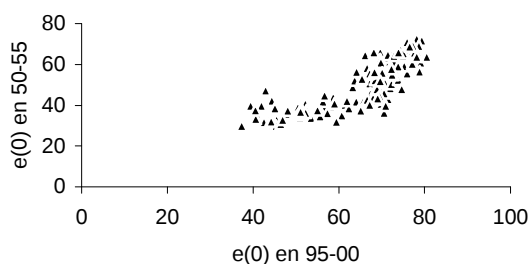
La première de ces figures se propose de localiser différents pays en fonction de leur espérance de vie en 1950-1955 et de leur gain d'espérance de vie entre 1950-1955 et 1995-2000³² (cf. figure 3.20). La deuxième de ces figures localise les pays en fonction de leur espérance de vie en 1950-1955 et en 1995-2000³³ (cf. figure 3.21).

Figure 3.20. Localisation en fonction de l'espérance de vie initiale et du gain d'espérance



(Source : United Nations (2001).)

Figure 3.21 Localisation en fonction des espérances de vie initiale et finale



(Source : United Nations (2001).)

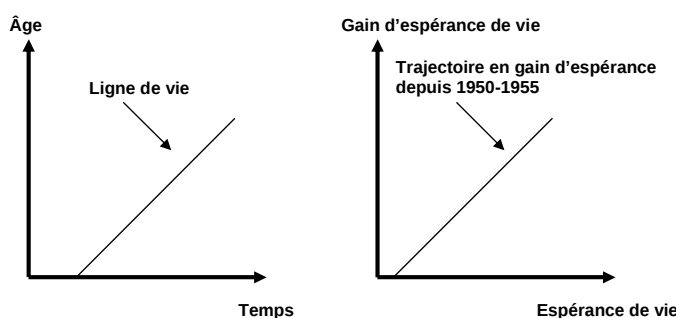
³² La figure originale (figure 3 en p. 7 de la publication de G. Caselli *et al.* (2002)) diffère par le fait que le gain d'espérance de vie portait sur la période allant de 1950-1955 à 1970-1975.

³³ La figure originale (la figure 14 en p. 19 de la publication de G. Caselli *et al.* (2002)) reprenait les espérances de vie en 1965 sur l'axe des abscisses et en 1995 sur l'axe des ordonnées. Nous avons mis en abscisses les données les plus récentes et non les plus anciennes ; on comprendra par la suite la raison de cette inversion.

Pour montrer que ces deux dernières figures méritent le qualificatif « de Lexis », nous allons en proposer... une troisième reprenant le même ensemble de données. Dans cette figure, l'espérance de vie jouera le rôle dévolu classiquement au temps, l'espérance de vie en 1950-1955, celui du moment de naissance et le gain d'espérance depuis 1950-1955, celui de l'âge (cf. figure 3.22).

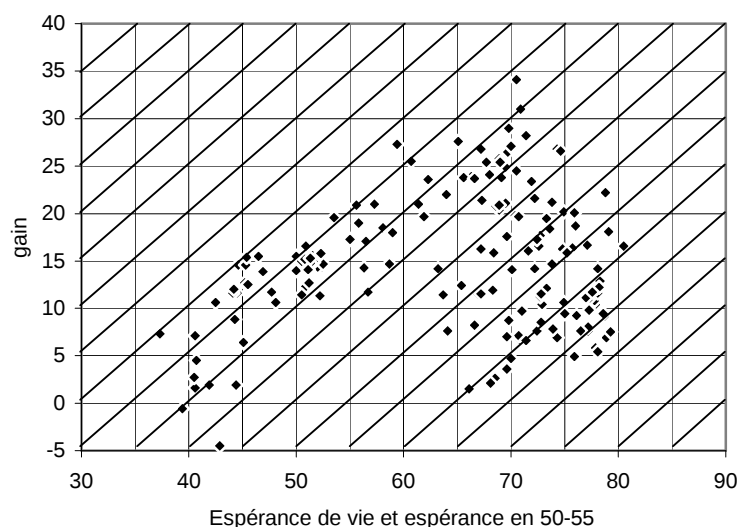
Sur ce diagramme, une trajectoire a été ajoutée ; elle montre l'évolution du gain d'espérance de vie en fonction de l'espérance de vie pour une espérance de vie en 1950-1955 donnée (expression équivalente à celle d'application dans le diagramme de Lexis classique : évolution de l'âge en fonction du temps pour un moment de naissance donné). Un seul point de cette trajectoire suffit pour la déterminer dans son ensemble. Autrement dit, ce type de trajectoire est comparable à une ligne de vie classique.

Figure 3.22. Un diagramme de Lexis en espérance de vie



Comment dans ce contexte définir une cohorte ? Une cohorte regroupera des trajectoires présentant une gamme de valeurs voisines pour leur espérance de vie en 1950-1955, par exemple, l'ensemble des trajectoires pour lesquelles cette espérance variait entre 45 et 50 ans. La figure 3.23 a été élaborée selon cette méthode de construction ; les diagonales y délimitent des cohortes quinquennales d'espérance de vie en 1950-1955.

Figure 3.23. Localisation en fonction de l'espérance de vie finale et du gain d'espérance



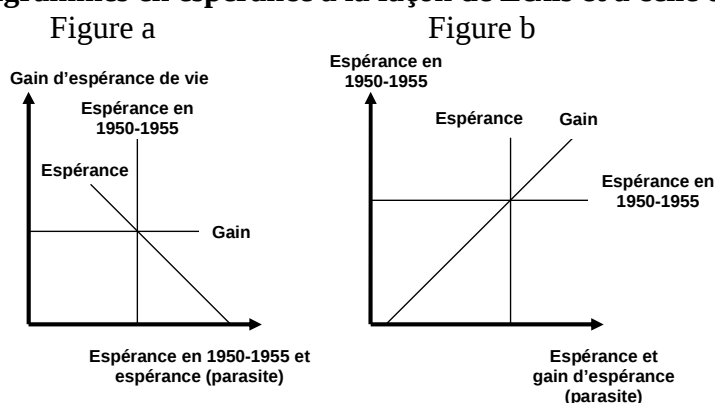
(Source : United Nations (2001).)

Quel est l'évènement matérialisé par les points de la figure 3.23 ? En fait, ces points permettent de connaître pour chaque pays à la fois l'espérance de vie en 1995-2000 en dessinant une verticale jusqu'à l'axe horizontal ; le gain d'espérance de vie entre 1950-1955 et 1995-2000 en traçant une horizontale jusqu'à l'axe vertical et finalement l'espérance de vie en 1950-1955 en traçant une diagonale jusqu'à l'horizontale en face d'un gain nul. Un point représente donc la période 1995-2000 pour une unité d'observation³⁴.

Attention, l'axe horizontal est bien dévolu à l'espérance de vie (et non à l'espérance de vie en 1995-2000) ; l'espérance de vie joue le rôle de variable « absolue » telle que définie sous le tableau 3.1. L'on pourrait être tenté d'en rester à la dénomination « espérance de vie en 1995-2000 » pour cet axe, mais cela équivaldrait, dans un diagramme de Lexis classique, à baptiser l'axe horizontal « moment du mariage » si les points sur le diagramme devaient représenter cet évènement.

Insistons, la situation décrite par la figure 3.23 est comparable à un diagramme de Lexis classique où les points représenteraient, par exemple, le décès des individus ou leur premier mariage : par une verticale, on connaîtra la date de l'évènement ; par une horizontale, leur âge à l'évènement et par une oblique, leur date de naissance. À notre avis, ceci suggère que le graphique 3.23 est un diagramme de Lexis, illustrant une analyse avec un cadre de référence en espérance de vie.

Figure 3.24. Diagrammes en espérance à la façon de Lexis et à celle de Becker



Qu'en est-il des figures 3.20 et 21 ? Il s'agit aussi de diagrammes de Lexis. En fait, la figure 3.20, correspond à un diagramme de Lexis tel que présenté au point 2.3.4.1, avec des trajectoires (et donc des cohortes) en position verticale et une coordonnée parasite (la valeur de l'espérance de vie) se matérialisant par des obliques décroissantes (cf. figure 3.24.a). Pour sa part, la figure 3.21 peut s'assimiler à un diagramme de Lexis tel que présenté au point 2.3.4.2, avec des trajectoires (et donc des cohortes) en position horizontale et une coordonnée parasite (le gain d'espérance de vie) se matérialisant par des obliques croissantes (cf. figure 3.24.b).

En fait, tout graphique mettant en relation des données individuelles d'une même variable à deux moments différents constitue un diagramme de Lexis, qui éventuellement s'ignore. Pour s'en rendre compte, il suffit de faire émerger, moyennant quelques calculs intermédiaires, la variable absolue (qui remplacera le temps du diagramme de Lexis classique), la valeur de cette variable à l'évènement-origine (qui remplacera le moment de naissance) et le gain en termes de cette variable par rapport à sa valeur à l'évènement-origine (qui remplacera l'âge).

³⁴ Pour l'interprétation des trajectoires forcées et des points, cf. annexe 3.9 ou le point 3.2.4.

Le système de repérage ainsi constitué au départ de deux coordonnées respectera les conditions de l'établissement d'un diagramme de Lexis. Finalement, en cas de données individuelles, et non plus collectives, l'exigence des trois coordonnées se réduit : il suffit que deux soient disponibles pour reconstituer l'ensemble des trois qui respectera les conditions précisées plus haut.

3.3.4. Conclusions

Le point 3.3 a été consacré à l'utilisation de diagrammes de Lexis avec des données inhabituelles. Par défaut d'autres possibilités, la généralisation s'est limitée à la représentation d'événements classés en fonction de variables certes non temporelles, mais dont les variations prises en compte étaient, elles, bien temporelles : le poids ou la taille des nourrissons ont été enregistrés au moment de la naissance et au moment de la fin d'observation, avec une variation mesurée entre ces deux moments.

Les diagrammes proposés portent sur l'évènement que constitue la sortie d'observation, le seul avec la naissance pour lequel les données nécessaires sont connues. Si le poids ou la taille étaient également enregistrés à d'autres événements (comme l'acquisition d'une capacité, la contraction d'une maladie, la perte de la première dent, le premier anniversaire, la perte du premier hochet...), des diagrammes construits selon les mêmes principes pourraient en montrer la distribution, ce qui pourrait en retour augmenter sensiblement la portée pratique de l'ouverture du diagramme de Lexis au monde des variables non strictement temporelles.

Les diagrammes en taille et en poids proposés ont mis en évidence des effets de la variation de la variable absolue (désignés par « effets d'âge » dans le diagramme classique) ainsi que, surtout dans le cas de la taille, des effets de cohorte (sans changement d'appellation, donc, par rapport à la situation classique).

Les premiers effets montrent simplement que, au sein des cohortes, les événements se dispersent autour d'une valeur centrale ; les seconds révèlent pour leur part que les cohortes adoptent, selon leur état initial, des comportements en matière de croissance nettement mieux différenciés en matière de taille qu'en matière de poids.

Dans un dernier point, nous avons montré que certaines figures sont en fait des diagrammes de Lexis. Il en va ainsi chaque fois qu'un graphique vise à mettre en relation les valeurs d'une même variable à deux dates différentes. Dans un tel contexte, les conditions mathématiques de l'établissement d'un diagramme de Lexis sont réunies (trois variables formant une combinaison linéaire ; une des trois invariable pour chaque unité d'observation).

3.4. Conclusions du chapitre 3

Le troisième chapitre avait pour objectif de généraliser et la construction et l'utilisation du diagramme de Lexis en lui faisant quitter le giron exclusif des variables temporelles. Cet objectif peut être considéré comme atteint.

Initialement conçu pour figurer des variations temporelles de variables temporelles, il a été montré que le diagramme de Lexis pouvait très bien s'adapter à un environnement constitué de variations de type quelconque d'une variable de nature quelconque. En définitive, pour construire une figure ayant les caractéristiques d'un diagramme de Lexis, il suffit de disposer de la mesure d'une même variable en deux circonstances différentes. La variable en cause joue le rôle classiquement dévolu au temps ; la valeur de cette variable à l'une des circonstances, celui du moment de naissance et sa variation entre les deux circonstances, celui de l'âge, l'évènement étudié étant la « deuxième » circonstance.

Cette généralisation du diagramme de Lexis a rejaili sur d'autres notions. Ainsi, un évènement peut être considéré, selon sa définition la plus générale, comme une circonstance pour laquelle les coordonnées de référence ont été enregistrées. Dans la foulée, il a aussi été montré que, sur un diagramme de Lexis, il convenait de distinguer la trajectoire et le trajet d'un individu : si la trajectoire regroupe tous les points du graphique qu'il est susceptible de fréquenter vu la valeur de la variable absolue à l'évènement-origine, le trajet se limite aux points effectivement fréquentés. Selon les circonstances, le trajet sera un simple extrait de la trajectoire ou au contraire une succession de points disjoints et/ou une suite d'allers-retours sur la trajectoire.

Au-delà de la simple construction d'un diagramme de Lexis avec variables non temporelles, la généralisation a aussi porté sur son utilisation effective. À l'aide d'exemples portant sur la taille et le poids des nourrissons, nous avons montré l'intérêt de diagrammes de Lexis pour montrer la distribution d'évènements même en cas de cadre de référence non temporel. L'équivalent d'effets d'âge et de cohorte a pu être mis en évidence. Par ailleurs, il a aussi été montré que certains graphiques sont en fait des diagrammes de Lexis qui s'ignorent.

Bref, ce chapitre a montré non seulement qu'il était possible d'élaborer sur le plan théorique des diagrammes de Lexis sur la base de variables de nature quelconque, mais en plus que de telles constructions pouvaient se révéler utiles pour représenter d'une manière originale de grandes masses de données et d'y faire apparaître des régularités potentiellement significatives dans l'analyse des phénomènes.

Ces points étant bien établis, le chapitre suivant abordera la généralisation des techniques d'analyse des démographes, à savoir, selon l'appellation proposée dans le premier chapitre, la cohortisation.