

CHAPITRE III

ÉTABLISSEMENT
DES HISTOIRES DE COHORTES

Le chapitre précédent a fixé, en dehors de toutes contingences, les types de description que l'on doit donner des différents phénomènes démographiques. Nous allons examiner maintenant comment utiliser les diverses sortes d'observations disponibles pour établir de telles histoires statistiques.

Débutons par quelques considérations sur les modes d'observation en démographie.

1 – LES MODES D'OBSERVATION DES COHORTES

D'après ce que nous venons de voir, chaque phénomène démographique implique, pour son étude, la constitution d'une cohorte spécifique. Cette cohorte doit être en observation pendant l'intervalle de temps $[0, \omega]$ durant lequel le phénomène se manifeste. Ainsi, dans l'exemple de l'étude de la nuptialité des femmes célibataires, l'origine 0 se place généralement au moment où les filles observées atteignent leur 15ème anniversaire et l'extrémité ω , lorsqu'elles atteignent 50 ans ; dans l'exemple de la divortialité l'origine est au mariage et l'extrémité environ 35 ans après. Si l'origine est toujours bien fixée, l'extrémité est plus imprécise et ne correspond jamais à la cessation véritable du phénomène, mais à une atténuation suffisamment importante de ces manifestations pour que l'on puisse négliger ce qui se passe au delà.

Pratiquement, les observations concernant une de ces cohortes peuvent être obtenues

a) soit en la suivant depuis son origine 0 et en enregistrant progressivement les événements la concernant (à la fois les événements de la catégorie étudiée et les événements perturbateurs).

b) soit en obtenant des membres de la cohorte qui dépassent l'extrémité ω , et lorsqu'il l'ont dépassée, les données sur les événements de la catégorie étudiée.

Dans le premier cas, les événements sont enregistrés au fur et à mesure qu'ils se produisent ; dans le second cas, ils le sont après que le phénomène s'est complètement manifesté.

Dans le premier cas les événements observés concernent tous les membres de la cohorte ; dans le second ils ne concernent que les membres encore présents (ceux qui ont survécu jusqu'à ω et qui n'ont pas émigré avant cette date).

La figure 21 symbolise ces deux modes d'observation. On dit, avec le premier, que l'on fait de l'observation continue (figure 21a) et avec le second que l'on fait de l'observation rétrospective (figure 21b)⁽¹⁾.

(1) Parallèlement à cette classification, qui concerne la façon dont les observations s'effectuent dans le temps, on introduit encore la notion d'observation suivie (Louis Henry), qui peut se faire également selon le mode continu ou le mode rétrospectif, mais dans laquelle les événements concernant un même individu sont rapprochés les uns des autres de manière à ce que l'on puisse tracer tout ou partie de la biographie de cet individu. Avec des observations suivies de nouvelles possibilités d'analyse sont offertes que nous n'examinerons pas dans cet ouvrage. Par ailleurs, pour certaines mesures de nature très classique, qui ne semblent pas requérir l'observation suivie, cette dernière est finalement indispensable ; ainsi, les études de mortalité par classes sociales qui exigent que les déclarations d'appartenance sociale concernant un individu soient les mêmes au moment de sa prise en observation et au moment de son décès : seule l'observation suivie peut assurer cette identité.

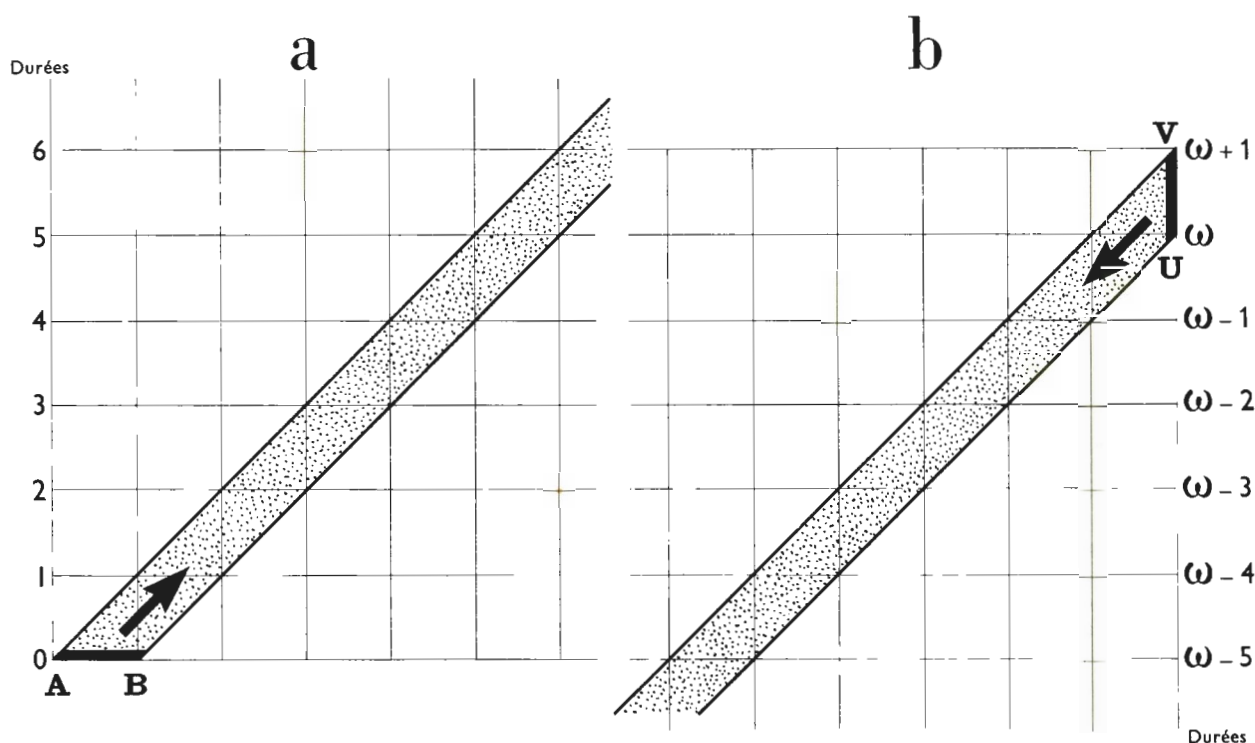


Figure 21 - Les deux modes d'observation, a) observation continue ; b) observation rétrospective.

L'observation rétrospective

D'un certain point de vue, c'est l'observation rétrospective qui permet d'aboutir le plus simplement aux histoires de cohortes dans les formes que nous venons de donner.

En effet, en n'interrogeant que des personnes ayant nécessairement échappé à certains événements perturbateurs⁽¹⁾, ceux-ci se trouvent par là même éliminés et, généralement, on atteint d'emblée les phénomènes à l'état pur. C'est par exemple de cette façon que l'on a obtenu les données sur la fécondité des générations féminines 1900-1909 en Grande-Bretagne (cf. tableau 18). On peut encore faire ainsi pour établir une table quelconque, par exemple une table de nuptialité féminine : dans ce dernier cas, on interroge les femmes atteignant 50 ans (ou celles formant un groupe d'âges au-dessus de 50 ans, 50-54 ans par exemple), pour savoir

- si elles ont conclu ou non un premier mariage
- en cas de premier mariage la date de celui-ci.

Les réponses fournissent directement la série des premiers mariages de la table de nuptialité et par conséquent les deux autres séries également.

Cette simplicité apparente de l'observation rétrospective ne doit pas faire oublier ses limitations et ses inconvénients :

(1) C'est ce qui a lieu par la force des choses : il faut que la personne interrogée soit présente, donc qu'elle ait échappé à la mort et n'ait pas émigré, catégories importantes d'événements perturbateurs. Mais, on peut concevoir que l'on atteigne, lors d'enquêtes rétrospectives, des personnes ayant vécu certains événements perturbateurs par rapport aux événements étudiés ; ainsi, dans une enquête de fécondité sous forme rétrospective on peut interroger des femmes devenues veuves ou ayant divorcé.

- elle ne saurait s'appliquer à l'étude de la mortalité, le propre de ce phénomène étant de faire disparaître les personnes qui sont objets de l'étude... On peut alors songer interroger des témoins de l'événement (parents, amis de la personne décédée), mais ce procédé est source d'imprécisions et d'omissions.

- l'observation rétrospective fait appel à la mémoire des enquêtés sauf si les événements cherchés ont laissé une trace écrite (livrets en possession des enquêtés) ; en l'absence de documents écrits, cet appel à la mémoire est source d'imprécisions et d'omissions.

- enfin, en n'interrogeant que des personnes ayant échappé nécessairement à certains événements perturbateurs, on opère des sélections, au sens statistique, dans les cohortes étudiées. En effet, les membres de la cohorte exclus de l'enquête parce qu'éliminés par la mort ou l'émigration, n'ont pas eu, en général, le même comportement vis-à-vis du phénomène étudié que les autres membres de la cohorte ; l'histoire que l'on écrit à partir des événements vécus par ces derniers, n'est donc pas identique à l'histoire de la cohorte entière. Donnons un exemple montrant de quelle façon un tel biais peut s'introduire : supposons que nous interrogeons les couples subsistants d'une promotion de mariages, à l'issue de 30 ans de mariage, délai nécessaire pour que toutes les naissances se soient produites ; les couples interrogés, puisqu'ils ont échappé à la mortalité, sont formés d'une plus grande proportion de couples mariés jeunes que n'en comprenait la promotion lors de sa constitution et, du même coup, la fécondité de la promotion s'en trouve surestimée, les couples mariés jeunes étant, pour différentes raisons, plus féconds que les autres(1).

Nous ne nous étendrons pas davantage sur cette forme d'observation, utilisée surtout lorsqu'on se limite à des ensembles d'assez faible dimension ; il est en effet difficile de la pratiquer à grande échelle en raison du soin particulier qui doit entourer la collecte des données auprès de chaque individu. En définitive, les problèmes qui se posent concernent plus les diverses phases de la collecte que l'élaboration des histoires statistiques découlant des observations recueillies.

L'observation continue

C'est encore la source principale de nos informations sur les faits démographiques. Il s'agit, par ailleurs, d'une forme d'observation très ancienne, antérieure à la démographie elle-même puisque pratiquée dès lors qu'il y a, dans un pays, enregistrement des faits d'état-civil ; c'est dire que l'observation continue ne concerne pas que les faits contemporains, mais aussi la masse considérable des faits passés ayant donné lieu à enregistrement, et qu'elle constitue la base la plus intéressante de la démographie historique(2).

Adaptée à l'étude de tous les phénomènes, l'observation continue est théoriquement la meilleure puisque, à la différence de l'observation rétrospective, elle ne néglige aucun des individus des cohortes étudiées ; en fait, elle présente elle aussi certaines limitations.

Pratiquée à l'échelle d'un pays entier, elle suppose, au niveau de la collecte, une bonne organisation administrative et une bonne discipline de la population, circonstances qui se trouvent réunies dans la plupart des pays développés ; mais, de toute façon, on ne saurait avec une opération aussi généralisée recueillir beaucoup d'informations entourant les événements démographiques eux-mêmes ; on est d'ailleurs prisonnier du système de relevés employé, système qui ne peut qu'évoluer lentement et dans des limites étroites. Enfin, au niveau de l'exploitation statistique, la lourdeur du travail est source d'erreur et limite l'étendue des résultats publiés, dont le choix résulte d'ailleurs plus souvent du maintien d'habitudes que de décisions correctement motivées, ce qui est domageable pour l'analyse.

(1) On peut sans doute, dans le cas présent, supprimer cette source de biais en étudiant la fécondité selon l'âge au mariage de la femme et en reconstituant éventuellement l'histoire de la promotion sur la base de sa structure, (connue ou estimée), selon l'âge de la femme au moment du mariage ; mais, il subsiste une autre source de biais, beaucoup plus générale et impossible à éliminer : celle provenant de la corrélation pouvant exister, toute question d'âge mise à part, entre la fécondité et la mortalité, c'est-à-dire entre le phénomène étudié et le ou les phénomènes perturbateurs.

(2) Base d'autant plus intéressante que la forme des renseignements recueillis permet l'observation suivie.

Pratiquée sur des échantillons restreints, spécialement choisis à cet effet, l'observation continue pose le difficile problème du maintien en observation des personnes considérées : l'expérience prouve qu'il faut toujours compter avec un certain pourcentage d'individus dont on perd la trace, ce qui nuit plus ou moins à la valeur des renseignements recueillis.

Dans la pratique, les conditions d'une parfaite observation continue sont rarement réunies ; on dispose plutôt d'informations plus ou moins cohérentes, plus ou moins complètes, pouvant s'ordonner à la manière des résultats d'une observation continue. Un grand travail d'élaboration est laissé à l'analyste pour transformer ces données brutes, qui pèchent toujours par quelques insuffisances, en véritables histoires de cohortes.

Nous aurons ainsi à examiner dans la suite, tout d'abord la question de l'utilisation des données d'une observation continue parfaite pour aboutir aux descriptions ayant fait l'objet du chapitre II, puis ensuite le problème, assez vaste et complexe, de la transformation et du rapprochement des données disponibles courantes, pour parvenir aux mêmes descriptions.

L'observation instantanée

Observation continue et observation rétrospective se déroulent dans le temps : il en est ainsi parce que ces observations concernent des événements démographiques qui se manifestent à des instants précis, et qui ne sauraient par conséquent être appréhendés statistiquement que sur des intervalles de temps.

Il existe, par contre, une autre forme d'observation en démographie, celle qui concerne les états de population et qui est de l'observation instantanée : il s'agit cette fois de saisir à un moment donné des situations en perpétuel changement, ce qui implique une simultanéité parfaite des opérations nécessaires pour connaître ces situations. Dans le diagramme de Lexis, une telle opération conduit à mesurer des flux.

L'exemple le plus commun d'observation instantanée est le recensement.

Nous serons ici très brefs sur cet autre aspect de la description démographique, puisqu'il fera l'objet du chapitre suivant. Remarquons seulement que par le jeu des naissances et des décès, il existe à tout moment des ensembles d'individus ou populations, dans lesquels on peut distinguer maintes catégories, en raison notamment des phénomènes démographiques déjà vécus : populations selon l'âge, l'état matrimonial, l'ancienneté du mariage, le degré d'instruction, etc...

Une observation continue parfaite est à même de nous renseigner sur ces caractéristiques d'état, sans qu'il soit nécessaire pour cela de monter des opérations spéciales. Mais, précisément, c'est parce qu'il n'y a jamais d'observation continue parfaite qu'on a recours à des observations instantanées, pour se renseigner sur la situation des diverses cohortes qui existent au sein d'une population.

Finalement, dans la pratique, ces données de l'observation instantanée complètent heureusement celles d'une observation continue toujours un peu déficiente, pour nous aider à tracer l'histoire statistique des cohortes. A la limite même, c'est parfois par la seule connaissance de l'état d'une cohorte à des dates successives que l'on parvient à décrire, d'une façon nécessairement approximative, le déroulement d'un phénomène.

2 – UTILISATION D'UNE OBSERVATION CONTINUE PARFAITE

D'après le titre, nous nous plaçons ici dans des conditions idéales, donc le plus souvent théoriques. Ce sera néanmoins une première étape indispensable avant d'envisager la façon d'utiliser les types de données les plus courants.

La mortalité

En ordonnant convenablement les données d'une observation continue relative à une génération, on aboutit immédiatement à la table de mortalité de cette génération. C'est d'ailleurs de cette façon que nous avons présenté la table de mortalité (cf. page 17).

La situation se complique lorsque la génération étudiée est soumise à des migrations, ou encore lorsque l'on se propose d'établir une table de mortalité résultant de la disparition d'une cause ou d'un groupe de causes de décès : on est alors en présence du problème de l'élimination d'un ou de plusieurs phénomènes perturbateurs (l'émigration ou la mortalité pour telle ou telle cause). Comme on rencontre d'emblée cette question lorsque l'on se propose de bâtir une table de nuptialité, nous l'éluciderons à cette occasion, et nous montrerons ensuite les formes particulières que prend la solution générale retenue.

La nuptialité

Le tableau 22 (quatre premières colonnes), offre un exemple des résultats que peut donner une observation continue, pratiquée dans une génération, et en ce qui concerne la nuptialité des célibataires (ici, des femmes).

L'observation est supposée débiter à 15 ans ; en face, à chaque âge x on a noté

l'effectif des célibataires à cet âge : C_x

les mariages de célibataires entre x et $x + 1$: $\mathcal{M}(x, x + 1)$

les décès de célibataires entre x et $x + 1$: $\mathcal{D}(x, x + 1)$

Nous avons pris soin d'employer des notations ne prêtant pas à confusion avec les quantités analogues des tables de nuptialité et de mortalité.

C'est parce qu'il y a possibilité de sortir du célibat de deux façons, par mariage ou par décès, que la construction d'une table de nuptialité, à partir des données de l'observation continue, n'est pas immédiate.

Nous avons analysé précédemment (cf. page 33) de quelle façon se produisent les $\mathcal{M}(x, x + 1)$ mariages effectifs entre les anniversaires x et $x + 1$, et aussi les $\mathcal{D}(x, x + 1)$ décès. Pour cela nous avons introduit

- la probabilité de se marier en l'absence de mortalité, n_x
- la probabilité de décéder en l'absence de nuptialité, q_x

Sous réserve que ces deux probabilités soient indépendantes, et au prix d'une approximation dans les calculs, nous avons établi les relations suivantes

$$e_x \left(1 - \frac{q_x}{2}\right) n_x = \pi(x, x+1)$$

$$e_x \left(1 - \frac{n_x}{2}\right) q_x = \omega(x, x+1)$$

Tableau 22
Observation continue dans une génération féminine et calcul des quotients de nuptialité

Age x	e_x	$\pi(x, x+1)$	$\omega(x, x+1)$	$e_x - \frac{\omega(x, x+1)}{2}$	$\frac{n_x}{(p. 1000)}$	Age x	e_x	$\pi(x, x+1)$	$\omega(x, x+1)$	$e_x - \frac{\omega(x, x+1)}{2}$	$\frac{n_x}{(p. 1000)}$
15	300.000	1.600	2.045	298.978	5,4	35	34.729	1.321	341	34.559	38,2
16	296.355	5.152	2.118	295.296	17,4	36	33.067	1.180	334	32.900	35,9
17	289.085	11.407	2.197	287.987	39,6	37	31.553	965	321	31.393	30,7
18	275.481	20.945	2.185	274.389	76,3	38	30.267	844	317	30.109	28,0
19	252.351	27.454	2.050	251.326	109,2	39	29.106	696	305	28.954	24,0
20	222.847	30.157	1.940	221.877	135,9	40	28.105	588	299	27.956	21,0
21	190.750	30.313	1.630	189.935	159,6	41	27.218	517	292	27.072	19,1
22	158.807	26.519	1.331	158.142	167,7	42	26.409	475	288	26.265	18,1
23	130.957	20.987	1.096	130.409	160,9	43	25.646	437	283	25.505	17,1
24	108.874	16.453	921	108.414	151,8	44	24.926	375	278	24.787	15,1
25	91.500	12.692	806	91.097	139,3	45	24.273	314	282	24.132	13,0
26	78.002	9.656	702	77.651	124,4	46	23.677	260	284	23.535	11,0
27	67.644	7.414	613	67.338	110,1	47	23.133	251	286	22.990	10,9
28	59.617	5.696	544	59.345	96,0	48	22.596	226	292	22.450	10,1
29	53.377	4.386	489	53.133	82,5	49	22.078	197	293	21.932	9,0
30	48.502	3.527	448	48.278	73,1	50	21.588				
31	44.527	2.701	421	44.317	60,9						
32	41.405	2.133	395	41.208	51,8						
33	38.877	1.842	372	38.691	47,6						
34	36.663	1.580	354	36.486	43,3						

C'est en utilisant ces relations, qui lient les données d'observation à certains éléments de la table cherchée (les quotients), que nous parviendrons à établir cette dernière

Nous avons là un système de 2 équations à 2 inconnues (n_x et q_x) qui conduit à une équation du second degré en n_x admettant une solution.

Mais nous pouvons obtenir une solution approchée de façon beaucoup plus simple. La première équation donne en effet

$$n_x = \frac{\pi(x, x+1)}{e_x - \frac{e_x q_x}{2}}$$

En assimilant $\frac{e_x q_x}{2}$ à $\frac{\omega(x, x+1)}{2}$ (ce faisant on commet une erreur par défaut égale à $\frac{e_x n_x q_x}{4}$ qui n'atteint jamais 10 % de la valeur de $\frac{e_x q_x}{2}$), on trouve

$$n_x = \frac{\pi(x, x+1)}{e_x - \frac{\omega(x, x+1)}{2}} \quad (F)$$

Ainsi, la formule donnant n_x , qui s'écrirait sans événements perturbateurs

$$\frac{\pi(x, x+1)}{e_x}$$

se trouve modifiée du fait de ces événements (les décès de célibataires), par l'adjonction du terme

$$- \frac{\mathcal{Q}(x, x+1)}{2}$$

au dénominateur.

En fait, la valeur de cette correction est souvent très faible (avec les données du tableau 22, elle est toujours inférieure à 6 ‰, bien que la mortalité dans cette génération soit élevée. C'est sensiblement celle définie par le tableau 4), et dans le cas d'une basse mortalité on peut la négliger.

C'est à l'aide de la formule (F) que l'on a déterminé les quotients du tableau 22. La table de nuptialité qui en découle se calcule aisément.

On tire pareillement du système d'équations initial,

$$q_x = \frac{\mathcal{Q}(x, x+1)}{e_x - \frac{e_x n_x}{2}}$$

En assimilant cette fois $\frac{e_x n_x}{2}$ à $\frac{\mathcal{N}(x, x+1)}{2}$ (on commet la même erreur par défaut, $\frac{e_x n_x q_x}{4}$ que précédemment), on trouve

$$q_x = \frac{\mathcal{Q}(x, x+1)}{e_x - \frac{\mathcal{N}(x, x+1)}{2}} \quad (F')$$

On notera le caractère symétrique des formules (F) et (F')⁽¹⁾, qui correspond aux rôles symétriques joués par les deux phénomènes en cause.

Une formule générale

Les formules (F) et (F') qui viennent d'être établies, ont une portée extrêmement générale : ce sont elles, ou d'autres adaptées aux divers cas particuliers mais analogues dans la forme, qui permettent de calculer, à partir des données d'une observation continue parfaite, les quotients nécessaires à l'établissement d'une table (au sens direct).

Donnons quelques exemples.

Faisons d'abord retour à la mortalité, en envisageant les cas déjà signalés où les événements perturbateurs s'interposent.

- si nous disposons, dans une génération sujette à émigration, des quantités

(1) Ces formules peuvent être établies par un raisonnement direct, sans faire appel au système d'équations en n_x et q_x (cf. par exemple Louis Henry, *Leçons d'Analyse démographique*, page 47). Notons qu'étant donné les approximations dont découlent ces formules, on n'a pas, avec les valeurs n_x et q_x obtenues, l'égalité

$$\frac{\mathcal{N}(x, x+1) + \mathcal{Q}(x, x+1)}{e_x} = n_x + q_x - n_x q_x$$

qui exprime le théorème des probabilités totales (il s'en faut d'ailleurs de très peu) ; par contre, ce théorème est vérifié avec les solutions du système d'équations en n_x et q_x .

$\mathfrak{S}_x$, effectif à l'âge x des personnes en observation

$\mathcal{O}(x, x + 1)$, nombre de décès entre les âges x et $x + 1$

$\mathcal{E}(x, x + 1)$, nombre d'émigrés entre les âges x et $x + 1$,

le quotient q_x est donné par la formule

$$q_x = \frac{\mathcal{O}(x, x + 1)}{\mathfrak{S}_x - \frac{\mathcal{E}(x, x + 1)}{2}} \quad (F'')$$

S'il y avait au contraire l'immigration $\mathcal{I}(x, x + 1)$, on aurait

$$q_x = \frac{\mathcal{O}(x, x + 1)}{\mathfrak{S}_x + \frac{\mathcal{I}(x, x + 1)}{2}}$$

D'une manière plus générale, on peut se borner à la formule (F'') dans laquelle $\mathcal{E}(x, x + 1)$ représente l'émigration nette ; il s'agit donc alors d'un nombre algébrique

- si nous voulons étudier, dans une génération fermée par exemple, la mortalité pour un groupe de causes noté 1, on aura, en notant 2 l'ensemble des autres causes.

$$q_{1,x} = \frac{\mathcal{O}_1(x, x + 1)}{\mathfrak{S}_x - \frac{\mathcal{O}_2(x, x + 1)}{2}}$$

On peut déterminer ainsi ce que deviendrait la mortalité si l'on parvenait à supprimer le cancer. Voici un exemple de calcul s'appliquant actuellement à une génération féminine d'un pays européen. Dans cette génération, le total des décès entre 67 et 68 ans est égal à

$$\mathcal{O}(67, 68) = 5.456$$

par ailleurs

$$\mathfrak{S}_{67} = 207.870$$

Le quotient annuel de mortalité à 67 ans vaut donc

$$q_{67} = \frac{5.456}{207.870} = 26,2 \text{ ‰}$$

Mais, sur les 5456 décès, 1.146 sont imputables au cancer ; le quotient de mortalité en l'absence de cancer est donc

$$q_{1,67} = \frac{4.310}{207.870 - \frac{1.146}{2}} = 20,8 \text{ ‰}$$

La suite $\{q_{1,x}\}$ que l'on détermine ainsi, fournit une table de mortalité, puis l'espérance de vie à la naissance $e_{1,0}$ qui lui correspond ; en comparant cette dernière à l'espérance de vie de la table "toutes causes", on mesure le gain en durée de vie à attendre de la suppression du cancer. L. Henry a trouvé de cette façon, pour les générations contemporaines, un gain possible de 1,8 ans chez les hommes et 2,1 ans chez les femmes.

Il peut naturellement y avoir plusieurs catégories d'événements perturbateurs, ainsi dans une étude de nuptialité, décès et émigration de célibataires ; on écrit alors

$$n_x = \frac{\mathfrak{N}(x, x + 1)}{\mathfrak{C}_x - \frac{\mathcal{O}(x, x + 1)}{2} - \frac{\mathcal{E}(x, x + 1)}{2}}$$

Faute de pouvoir passer en revue tous les phénomènes, terminons par un exemple⁽¹⁾ qui marquera bien la généralité du procédé. Il s'agit (tableau 23), de données sur les remariages de veufs. Les quotients de nuptialité que l'on doit calculer résultent de la formule

$$n'_x = \frac{\mathcal{R}(x, x + 1)}{v_x - \frac{\mathcal{D}(x, x + 1)}{2}}$$

Tableau 23

Observation continue dans une cohorte de veufs à 40-49 ans et calcul des quotients de nuptialité

Durée du veuvage x	v_x	$\mathcal{R}(x, x + 1)$	$\mathcal{D}(x, x + 1)$	$v_x - \frac{\mathcal{D}(x, x + 1)}{2}$	n'_x (p. 1000)
0	500	150	10	495	303
1	340	80	8	336	238
2	252	40	8	248	161
3	204	20	6	201	100
4	178	10	6	175	57
5	162	...	...	...	...
...	...	...	...	...	...

La table de nuptialité s'en déduit comme n'importe quelle table définie par la suite des quotients.

La fécondité

L'utilisation des données d'une observation continue parfaite pour établir des tables de fécondité de rang, n'est qu'un cas particulier du problème général que nous venons de résoudre.

Il ne nous reste donc à étudier que la façon d'utiliser les résultats d'une telle forme d'observation pour établir la série des descendance $\{D_x\}$ ou des naissances $\{n(x, x + 1)\}$ (cf. tableaux 19 et 20), définie

- soit dans une promotion de mariages
- soit dans une génération

Posons et résolvons le problème tout d'abord dans le cas d'une promotion de mariages. Les données d'une observation continue parfaite prennent alors la forme des quatre premières colonnes du tableau 24 : à un anniversaire x il subsiste u_x unions parmi les u_0 (ici 4.431) qu'il y avait à l'origine, et entre x et $x + 1$ il se produit

$\mathcal{N}(x, x + 1)$ naissances

$\mathcal{O}(x, x + 1)$ dissolutions d'unions (veuvages et divorces)

Le problème est de dégager, à partir de ces observations, les manifestations de la fécondité à l'état pur.

(1) Cf. Roland Pressat. *Pratique de la Démographie. Trente sujets d'analyse*. Sujet n° 16.

Tableau 24

Observation continue parfaite dans une promotion de mariages et calcul des naissances en l'absence de dissolutions d'unions

Anniversaire du mariage x	u_x	$\pi(x, x + 1)$	$\omega(x, x + 1)$	$\frac{u_x + u_{x+1}}{2}$	u_x (p. 1.000)
0	4.431	1.526	65	4.398	347
1	4.366	1.151	67	4.332	266
2	4.299	905	70	4.264	212
3	4.229	744	74	4.192	177
4	4.155	607	76	4.117	147
5	4.079	537	77	4.040	133
6	4.002	448	77	3.963	113
7	3.925	381	76	3.887	98
8	3.849	296	74	3.812	78
9	3.775	272	73	3.738	73
10	3.702	245	71	3.666	67
11	3.631	208	70	3.596	58
12	3.561	175	70	3.526	50
13	3.491	157	69	3.456	45
14	3.422	133	68	3.388	39
15	3.354				...
...					...

Dans le tableau 24 ces manifestations sont troublées par l'existence d'une mortalité des unions (au sens large : veuvage et divorces), que l'on peut mesurer directement en rapprochant les u_x et les $\omega(x, x + 1)$ et en calculant les quotients q_x correspondants (cf. tableau 25).

Tableau 25

Table de mortalité des unions

Anniversaire du mariage x	u_x	$\omega(x, x + 1)$	q_x (p. 1.000)
0	4.431	65	15
1	4.366	67	15
2	4.299	70	16
3	4.229	74	17
4	4.155	76	18
5	4.079	77	19
6	4.002	77	19
7	3.925	76	19
8	3.849	74	19
9	3.775	73	19
...			

Quant à la fécondité à l'état pur, elle peut se définir par la série des naissances $\{N(x, x + 1)\}$ aux différentes durées de mariages correspondant à u_0 mariages non soumis à mortalité, ou encore par les fréquences

$$u_x = \frac{N(x, x + 1)}{u_0}$$

qui ont la signification d'un nombre moyen de naissances vivantes par an et par union.

Il nous reste à établir une équation liant des quantités résultant de l'observation à ce que nous cherchons, u_x par exemple. Pour cela nous ferons encore l'hypothèse que la fécondité et la mortalité des unions sont indépendantes ; c'est ainsi, par exemple, que nous donnerons aux u_x unions subsistantes à la durée de mariage x , la fécondité u_x définie en l'absence de désunions.

On trouve ainsi entre les anniversaires x et $x + 1$

$u_x(1 - q_x)$ u_x naissances dans les unions qui survivent en $x + 1$

$u_x \frac{q_x u_x}{2}$ naissances dans les unions qui sont détruites avant $x + 1$

soit au total

$$u_x \left(1 - \frac{q_x}{2}\right) u_x \text{ naissances effectives}$$

Comme ici

$$u_{x+1} = u_x(1 - q_x)$$

cette formule peut s'écrire

$$\frac{u_x + u_{x+1}}{2} \cdot u_x$$

Ces naissances effectives sont les $\mathcal{N}(x, x + 1)$ naissances du tableau 24 ; nous avons donc l'équation en u_x

$$\mathcal{N}(x, x + 1) = \frac{u_x + u_{x+1}}{2} \cdot u_x$$

et par conséquent

$$u_x = \frac{\mathcal{N}(x, x + 1)}{\frac{u_x + u_{x+1}}{2}} \quad (G)$$

formule que l'on peut encore écrire

$$u_x = \frac{\mathcal{N}(x, x + 1)}{u_x - \frac{\mathcal{D}(x, x + 1)}{2}}$$

si l'on veut conserver la structure des formules (F), (F'), et (F'').

Le raisonnement précédent vaut pour la détermination, à partir d'une observation continue parfaite, des descendance D_x ou des naissances $n(x, x + 1)$ dans une génération (cf. tableau 20).

On dispose alors, à un anniversaire x quelconque (compris entre 15 et 50 ans)

- des femmes encore en vie : $\mathcal{F}_x$
- des naissances entre x et $x + 1$ provenant de ces femmes : $\mathcal{N}(x, x + 1)$
- des décès de femmes entre x et $x + 1$: $\mathcal{D}(x, x + 1)$ (1)

~~Les quatre premières colonnes du tableau 26 rapportent un tel exemple d'observation.~~

La fécondité à l'état pur peut se caractériser cette fois par la série des naissances $\{N(x, x + 1)\}$ aux différents âges, correspondant à $\mathcal{F}_{15}$ femmes non soumises à mortalité, ou encore par les fréquences

(1) Pour ne pas multiplier les notations, nous reprenons certains symboles déjà employés, dans la mesure où ils ne prêtent pas à confusion.

$$f_x = \frac{N(x, x+1)}{\mathfrak{F}_{15}}$$

qui ont la signification d'un nombre moyen de naissances vivantes par an et par femme.

Avec la même hypothèse d'indépendance de la fécondité et de la mortalité on aboutit à

$$f_x = \frac{\mathfrak{N}(x, x+1)}{\frac{\mathfrak{F}_x + \mathfrak{F}_{x+1}}{2}} \quad (G')$$

qui peut encore s'écrire, si l'on veut conserver la structure des formules (F) (F') et (F''),

$$\frac{\mathfrak{N}(x, x+1)}{\mathfrak{F}_x - \frac{\mathcal{D}(x, x+1)}{2}}$$

En résumé, dans les deux situations que nous venons d'analyser, nous sommes parvenus à dégager la fécondité pure en calculant les indices

$$u_x \quad \text{et} \quad f_x$$

qui représentent des nombres moyens de naissances par an et par tête (union ou femme), en l'absence de mortalité. On passe de là immédiatement aux séries de naissances $\{n(x, x+1)\}$ et de descendance $\{D_x\}$ que nous avons présentées dans les tableaux 19 et 20.

Ces indices u_x et f_x se nomment des taux, notion fort utilisée en démographie.

Un taux se caractérise par le fait que les événements qui entrent dans son calcul sont rapportés à un effectif moyen et non à un effectif initial comme dans les quotients.

A vrai dire, cette distinction, si elle aide à fixer le langage, est au fond bien fragile : la référence à l'effectif moyen laisse place, en fait, à bien des possibilités et par conséquent à des indices de significations fort différentes.

Etant donné la place que tiennent les taux en démographie et qu'ils prendront dans la suite de notre exposé, nous allons leur accorder dès à présent une attention particulière(1).

Digression sur les taux en démographie

S'agissant de la localisation des événements que l'on considère, on peut, dans la pratique (et dans l'hypothèse d'un intervalle d'une année), rencontrer les trois possibilités de la figure 22.

- avec la figure 22-a, nous retrouvons la disposition des données du calcul de u_x et de f_x : $\mathfrak{N}(x, x+1)$ dans AA'BB', u_x (ou $\mathfrak{F}_x$) en AB et u_{x+1} (ou $\mathfrak{F}_{x+1}$) en A'B'. La moyenne arithmétique.

$$\frac{u_x + u_{x+1}}{2}$$

qui représente l'effectif moyen, est approximativement le flux en MN, qu'on peut assimiler à celui en A'B.

- avec la figure 22-b, l'intervalle d'âge ou de durée est mal défini : il varie de $[x-1, x]$ à $[x, x+1]$ selon que l'on se déplace en AA' ou en BB' ; on peut l'assimiler en moyenne à l'intervalle $\left[x - \frac{1}{2}, x + \frac{1}{2} \right]$. En revanche, la localisation dans le temps est parfaite.

(1) Toutefois, nous n'examinerons maintenant que les taux par âge ou par durée.

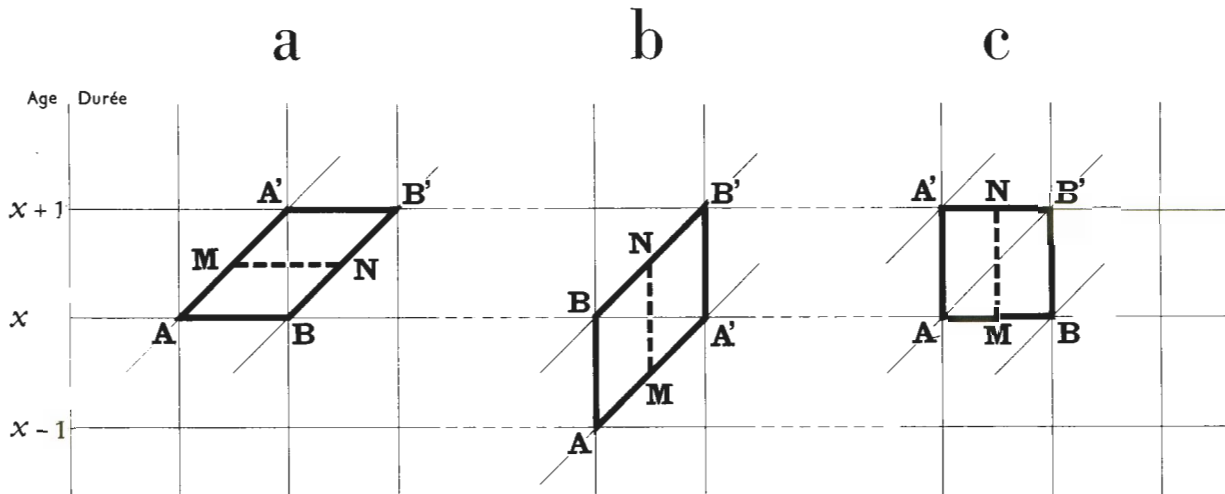


Figure 22 - Emplacement des éléments dans différentes sortes de taux.

- la figure 22-c est une configuration peu satisfaisante bien que fort utilisée (sans doute parce que intervalles de durées et de temps sont tous les deux très précisément délimités) : deux cohortes se trouvent mélangées et l'effectif moyen s'y définit mal.

Cependant, ce qui différencie le plus les taux, ce ne sont pas ces diverses possibilités de configuration, mais le choix de l'effectif moyen entrant dans le dénominateur. Nous allons illustrer cela en utilisant les données sur la nuptialité des célibataires figurant au tableau 22.

A un âge x , nous connaissons l'effectif C_x et aussi l'effectif de la génération S_x (cette donnée ne figure pas dans le tableau 22 parce qu'elle ne nous était pas utile) ; entre x et $x + 1$, il y a $\pi(x, x + 1)$ premiers mariages et $\varnothing(x, x + 1)$ décès de célibataires.

On peut envisager de calculer un taux pour cet intervalle $[x, x + 1]$,

- soit en prenant l'effectif moyen des célibataires :

$$t_x = \frac{\pi(x, x + 1)}{\frac{C_x + C_{x+1}}{2}} \quad (T)$$

- soit en prenant l'effectif moyen de la génération :

$$t'_x = \frac{\pi(x, x + 1)}{\frac{S_x + S_{x+1}}{2}} \quad (T')$$

L'idée qui préside au premier calcul est au fond celle qui est à la base du calcul des quotients : mesurer un risque, c'est-à-dire ici, rapporter des mariages à des mariages. Il s'ensuit qu'il existe une relation simple entre un taux ainsi compris et un quotient. Abandonnons provisoirement les indices pour alléger l'écriture. La formule (T) s'écrit encore

$$t = \frac{\pi}{C - \frac{\varnothing}{2} - \frac{\pi}{2}}$$

tandis que

$$n = \frac{\pi}{C - \frac{\varnothing}{2}} \quad \text{d'où} \quad C - \frac{\varnothing}{2} = \frac{\pi}{n}$$

En utilisant cette dernière relation il vient

$$t = \frac{\mathfrak{M}}{\frac{\mathfrak{M}}{n} - \frac{\mathfrak{M}}{2}} = \frac{1}{\frac{1}{n} - \frac{1}{2}}$$

Finalement

$$t = \frac{2n}{2 - n} \quad \text{et} \quad n = \frac{2t}{2 + t} \quad (1)$$

t est plus grand que n comme on pouvait le voir directement sur la formule de définition ; mais, dans la mesure où n est faible, les deux grandeurs sont très voisines. Ainsi, avec les données du tableau 22

$$n_{30} = 73,1 \text{ ‰} \quad \text{et} \quad t_{30} = \frac{2 \times 0,0731}{2 - 0,0731} = 75,9 \text{ ‰}.$$

La formule (T') donne un tout autre résultat qui revêt une toute autre signification. Essayons de rattacher le taux, ainsi compris, à un des éléments de la table de nuptialité correspondante.

Nous avons établi précédemment (cf. page 68) que

$$\mathfrak{M}(x, x + 1) = \mathfrak{C}_x \left(1 - \frac{q'_x}{2}\right) n_x$$

où q'_x représente le quotient de mortalité des seuls célibataires. Mais en faisant l'hypothèse d'indépendance de la mortalité et de la nuptialité, on trouve que $\mathfrak{C}_x$ vaut

$$\mathfrak{S}_{15} \frac{C_x}{C_{15}} \frac{S'_x}{S_{15}}$$

C_x et S'_x indiquant les survivants des tables de nuptialité et de mortalité des célibataires.

Par ailleurs

$$\frac{\mathfrak{S}_x + \mathfrak{S}_{x+1}}{2} = \frac{\mathfrak{S}_{15}}{2} \cdot \frac{S_x + S_{x+1}}{S_{15}}$$

S_x indiquant les survivants de la table de mortalité de l'ensemble de la génération.

En rassemblant ces résultats, et en se souvenant que $C_x n_x = m(x, x + 1)$ (mariages de la table de nuptialité) il vient

$$t'_x = \frac{m(x, x + 1)}{C_{15}} \frac{S'_{x+0,5}}{S_{x+0,5}}$$

Autrement dit

$$m(x, x + 1) = t'_x C_{15} \frac{S_{x+0,5}}{S'_{x+0,5}}$$

Ainsi, à un facteur près, le taux donne, dans ce second cas, les mariages de la table de nuptialité. Ce facteur,

(1) Ces formules sont valables dans les deux premiers cas de la figure 22 (22-a et 22-b). Elles sont vérifiées approximativement dans le troisième cas (22-c) si les deux générations en cause ont eu des effectifs initiaux et des histoires à peu près semblables ; la formule dans ce cas n'est utilisable qu'après avoir défini le mode de calcul d'un quotient dans un carré de Lexis, ce que nous ferons plus loin.

$$\frac{S_{x+0,5}}{S'_{x+0,5}}$$

est généralement un peu supérieur à l'unité, en raison de la surmortalité des célibataires : en identifiant purement et simplement le taux t'_x aux mariages $m(x, x + 1)$, on sous-estime un peu la nuptialité.

Dans notre exemple, nous avons (ces valeurs n'ont pas été rapportées dans le tableau 22 car elles n'étaient d'aucune utilité),

$$S_{30} = 262.817, S_{31} = 260.283$$

comme

$$\pi(30, 31) = 3.527$$

il en découle

$$t'_{30} = \frac{3.527}{261.550} = 13,5 \text{ ‰}$$

Le lecteur est invité à utiliser les quotients de nuptialité du tableau 22 pour amorcer la construction de la table de nuptialité correspondante : en prenant $C_{15} = 10.000$, il trouvera $m(30, 31) = 135$; ici il y a donc identité (à un multiple de 10 près) entre t'_{30} et $m(30, 31)$, la différence de mortalité entre célibataires et non célibataires étant insignifiante dans cet exemple.

Les résultats absolument généraux que nous venons de trouver appellent deux remarques :

- les taux de mortalité sont nécessairement des taux de la première catégorie ; on ne peut en effet rapporter les décès qu'aux vivants puisqu'il n'est pas habituel de donner les effectifs d'une génération en distinguant les vivants et les morts

- les taux u_x et f_x sont à rapprocher des taux de deuxième catégorie, les résultats que nous avons établis à leur sujet étant analogues à ceux concernant la forme (T') appliquée aux phénomènes représentables par des tables : ils ne sauraient d'ailleurs être de la première catégorie (forme (T)) ces derniers taux étant particuliers aux événements non renouvelables.

En conclusion les divers taux par âge ou durée que l'on peut calculer en démographie se distinguent essentiellement par le fait qu'ils peuvent contenir au dénominateur.

- soit les seules personnes de la cohorte n'ayant pas encore vécu le phénomène étudié (taux de première catégorie, forme (T)).

- soit l'ensemble des personnes de la cohorte, qu'elles aient ou non vécu le phénomène (taux de deuxième catégorie, forme (T')).

La première forme donne des quantités s'apparentant à des quotients et la seconde des quantités sensiblement identiques aux événements des tables au sens strict ou des tableaux de descendance de promotion ou de génération.

Il s'ensuit que la seconde forme est au moins aussi précieuse que la première puisqu'elle fournit de façon pour ainsi dire immédiate l'histoire statistique cherchée et que, de surcroît, elle exige une information moins détaillée.

Vues d'ensemble

Lorsque l'on dispose d'une observation continue parfaite, l'histoire statistique des diverses cohortes s'établit par l'intermédiaire du calcul de quotients ou de taux.

Lors d'un tel calcul, l'élimination des phénomènes perturbateurs est faite en posant l'hypothèse d'indépendance entre ces phénomènes et ceux que l'on étudie ; cette hypothèse n'est probablement jamais rigoureusement vérifiée et, dans certains cas, la réalité s'en écarte suffisamment pour limiter sensiblement la signification des résultats. Ainsi par exemple :

- la mortalité des célibataires est mal mesurée car ce que l'on atteint ne résulte pas seulement de la mortalité liée à la condition de célibataire, mais aussi de l'effet de sélection qui maintient dans le célibat certaines personnes présentant certains handicaps physiques, facteurs de plus haute mortalité.

- la mesure des résultats de l'élimination d'une cause ou d'un groupe de causes de mortalité est tout aussi hasardeuse si l'indépendance supposée de ces causes avec l'ensemble des autres causes n'est pas vérifiée avec suffisamment de précision.

Après cet examen de l'emploi des données d'une observation continue parfaite, nous allons étudier les problèmes posés par l'utilisation des statistiques démographiques existantes.

3 – UTILISATION DES STATISTIQUES DÉMOGRAPHIQUES COURANTES

Rappelons que nous avons toujours en vue l'établissement d'histoires de cohortes dans les formes qui ont été précisées au chapitre II ; nous venons de voir qu'avec une observation continue parfaite, c'est le calcul de quotients ou de taux qui permet de résoudre ce problème. Il en va souvent de même ici, mais la variété des situations concrètes obligera parfois à considérer diverses solutions de remplacement.

Les statistiques concernant les divers phénomènes démographiques peuvent revêtir maints aspects que l'on ne peut songer passer systématiquement en revue, dans un exposé de caractère général. Nous nous attacherons ici à l'examen des formes les plus habituelles, dans un pays à bonnes statistiques démographiques comme la France, et en n'examinant que les phénomènes les plus courants. Les analyses qui seront faites à cette occasion devraient permettre au lecteur de se tirer d'affaire, au prix d'initiatives peu importantes, dans la plupart des problèmes concrets⁽¹⁾.

Les informations qui seront à notre disposition, dans les divers cas que nous examinerons, seront toujours à rattacher à celles que donnerait une observation continue parfaite : c'est ce rapprochement qui nous aidera, de façon plus ou moins directe, à trouver les solutions des divers problèmes considérés.

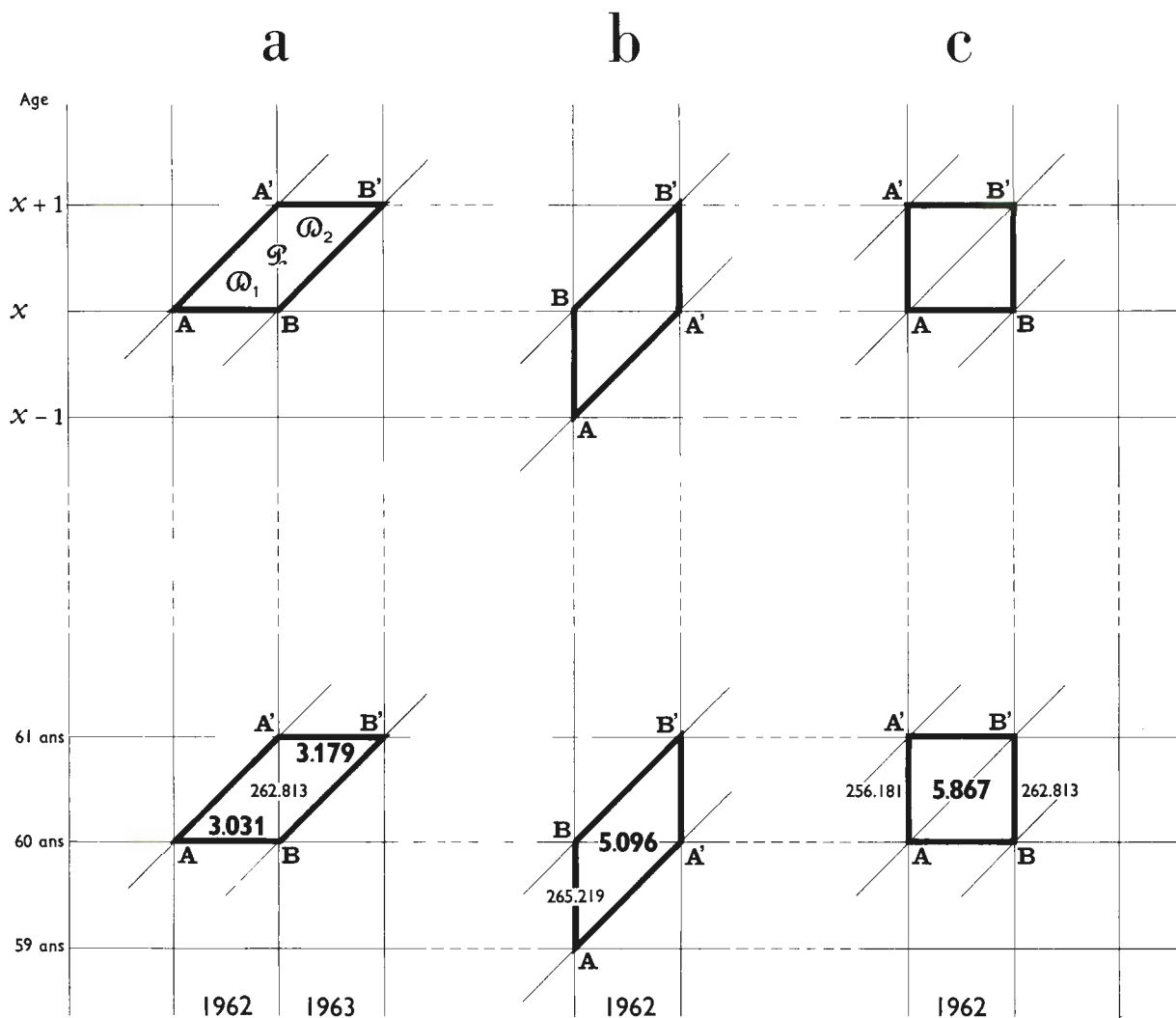
La mortalité

Posons comme remarque, de portée absolument générale, que les conditions de l'observation courante ne se prêtent jamais à la connaissance directe des flux au travers de segments horizontaux du diagramme de Lexis (sauf lorsqu'il s'agit des nouveau-nés, des nouveaux mariés, des nouveaux divorcés, ... d'une année) ; les recensements ou les diverses opérations qui en tiennent lieu ne peuvent donner que des flux au travers de segments verticaux.

(1) Nous ne soulèverons pas ici les questions posées par l'utilisation de données de qualité insuffisante.

Les circonstances concrètes les plus favorables pour l'étude de la mortalité sont celles où l'on peut combiner les statistiques d'état de la population avec les statistiques de décès par âge et par génération. On est dans le cas de figure 23-a et l'on a visiblement

$$q = \frac{\omega_1 + \omega_2}{x + \omega_1}$$



$$t_{60} = \frac{5.867}{\frac{256.181 + 262.813}{2}} = 22,6 \text{ ‰}$$

$$q_{60} = \frac{3.031 + 3.179}{262.813 + 3.031} = 23,4 \text{ ‰} \quad q_{59'} = q_{59,5} = \frac{5.096}{265.219} = 19,2 \text{ ‰} \quad q_{60} = \frac{2 \times 0,0226}{2,0226} = 22,3 \text{ ‰}$$

Figure 23 - Calcul d'un quotient de mortalité. Divers cas de figures. Exemples empruntés à la population masculine française.

Le seul phénomène perturbateur possible est la migration ; notant $\mathcal{E}_1$ et $\mathcal{E}_2$ la valeur de l'émigration nette dans les deux triangles de Lexis, la formule précédente se transforme en

$$\frac{\omega_1 + \omega_2}{\mathcal{Q} + \omega_1 + \mathcal{E}_1 - \frac{\mathcal{E}_1 + \mathcal{E}_2}{2}}$$

Or $\mathcal{E}_1 - \frac{\mathcal{E}_1 + \mathcal{E}_2}{2} \simeq 0$ si $\mathcal{E}_1 \simeq \mathcal{E}_2$ ce qui arrive si les courants de migration sont réguliers ; ainsi, dans ces conditions (assez souvent remplies), la première formule reste valable.

Une variante intéressante est fournie par la figure 23-b : elle n'exige que la connaissance des décès par génération et année civile. On trouve un quotient que nous avons déjà rencontré et noté alors

$$q_x,$$

pour rappeler que l'âge initial x' est un âge en années révolues ; ce quotient peut être assimilé à

$$q_{x+0,5}$$

et l'on peut construire la table de mortalité correspondante (notre figure correspond au calcul du quotient $q_{(x-1),}$, que l'on peut assimiler au quotient $q_{x-0,5}$) ; ce n'est toutefois pas l'usage et il convient plutôt de passer, par interpolation, de la série $\{q_{x'}\} = \{q_{x+0,5}\}$ à la série $\{q_x\}$.

Dans un pareil cas de figure, la correction pour migrations se fait selon la formule classique.

En présence d'informations occupant tout le carré ABA'B' de la figure 23-c on doit renoncer, de toute façon, à une imputation à une génération précise ; en suivant la suite des carrés ayant la droite AB' comme diagonale commune, on suit toujours (mais avec des lacunes), le même couple de générations. On peut alors calculer le taux t correspondant à chaque carré et passer de là au quotient q par la formule déjà établie (cf. page 76 et la note 1) :

$$q = \frac{2t}{2 + t}$$

Nous donnons dans la figure 23 des exemples de calculs⁽¹⁾.

Après ces cas qui nous rattachent encore assez étroitement à l'observation continue, nous devons évoquer l'éventualité de la seule connaissance, de loin en loin, de l'effectif de la génération : les circonstances peuvent être variées et se présentent surtout dans les pays à état civil insuffisant ou inexistant. Les populations recensées fourniront alors, après des traitements divers, quelques éléments de la série de survivants de la table que l'on cherche. On peut imaginer, par exemple, que l'on connaisse, dans un groupe de cinq générations, les effectifs $\mathcal{Q}_{5-9}$ et $\mathcal{Q}_{10-14}$: s'il n'y a pas de migrations, on assimilera ces effectifs à $\mathcal{S}_{7,5}$ et $\mathcal{S}_{12,5}$.

Signalons comme cas particulier de l'éventualité précédente, celui où nous possédons l'effectif d'une génération, ou d'un groupe de générations, à une certaine date et l'effectif des nouveau-nés correspondants (figure 24) ; ce peut être ainsi l'effectif en A'B' que l'on compare à celui en AB ou celui en A'C que l'on compare à celui en AC. Signalons que l'égalité

(1) Dans le cas de figure b, on peut se proposer d'effectuer la correction pour migration, d'autant qu'il y a eu une immigration particulièrement importante en France en 1962. Comme la population en BB' est de 262.813 (cf. BB' sur la figure c), l'immigration nette dans la génération au cours de 1962 est de $262.813 - (265.219 - 5.096) = 2.690$; d'où le nouveau quotient

$$q_{59} = q_{59,5} = \frac{5.036}{265.219 + 1345} = 19,1 \text{ \%}$$

Remarquons encore que les divers quotients que nous venons de calculer ne sauraient se comparer, car ils diffèrent soit par la période de calcul soit par l'intervalle d'âge.

flux en AC $\times p$ = flux en A'C,

qui correspond à ce dernier cas de figure, est d'ailleurs souvent utilisée en pays sous-développés pour évaluer le flux en AC (naissances) à partir de la donnée du flux en A'C et d'une estimation de p (1).

La nuptialité

L'examen des diverses statistiques courantes de nuptialité, va nous permettre de faire le tour complet des méthodes usuelles d'étude des phénomènes caractérisés par des événements non renouvelables.

En premier lieu, nous nous référons aux trois configurations déjà considérées ; nous les rappelons dans la figure 25.

Avec les données de la figure 25-a, analogues à celles de la figure 23-a, le quotient de nuptialité est donné par,

$$n = \frac{\pi_1 + \pi_2}{c + \pi_1}$$

les événements perturbateurs (décès de célibataires et éventuellement migrations) se trouvant éliminés de la formule de la même façon que pour la mortalité.

Les configurations b et c peuvent sans doute donner lieu à des calculs directs de quotients, mais comme la possibilité de calculer des taux (de la 2ème catégorie) s'offre aussi à nous (ce qui ne pouvait avoir lieu avec la mortalité), nous donnerons la préférence à cette seconde solution : nous parvenons en effet ainsi de façon plus directe au résultat cherché, en utilisant d'ailleurs une information moins détaillée, ce qui correspond à une réalité plus courante. Nous rappelons que du taux t'_x ainsi calculé, nous passons aux mariages $m(x, x + 1)$ par la formule

$$m(x, x + 1) = t'_x C_{15} \frac{S_{x+0,5}}{S'_{x+0,5}}$$

Dans la mesure où nous négligeons le facteur correctif (et faute d'informations il est généralement difficile d'en tenir compte), nous sous-estimons la nuptialité.

Enfin, une dernière possibilité s'offre à nous, qui ne se présentait pas lors de l'examen des statistiques de mortalité : c'est l'utilisation de la répartition aux divers anniversaires, de l'effectif de la génération en célibataires et non célibataires.

On conçoit, instinctivement, que cette répartition soit un reflet de la nuptialité. Nous allons établir rigoureusement le lien qui existe entre cette répartition et certains éléments de la table de nuptialité.

Nous avons établi précédemment qu'à un anniversaire x , l'effectif des célibataires c_x est donné par

$$c_x = S_{15} \cdot \frac{C_x}{C_{15}} \cdot \frac{S'_x}{S'_{15}}$$

(1) Cf. Roland Pressat. *Pratique de la démographie. Trente sujets d'analyse.* Sujet n° 20.

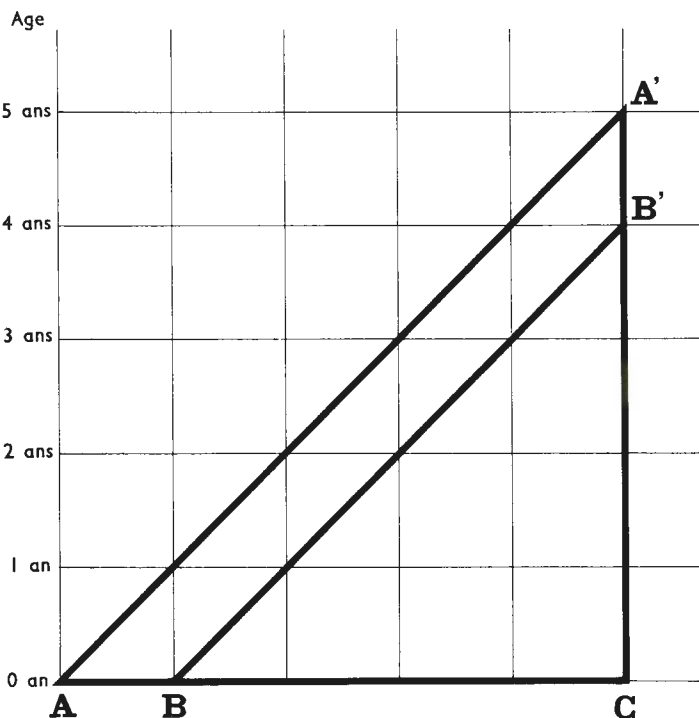
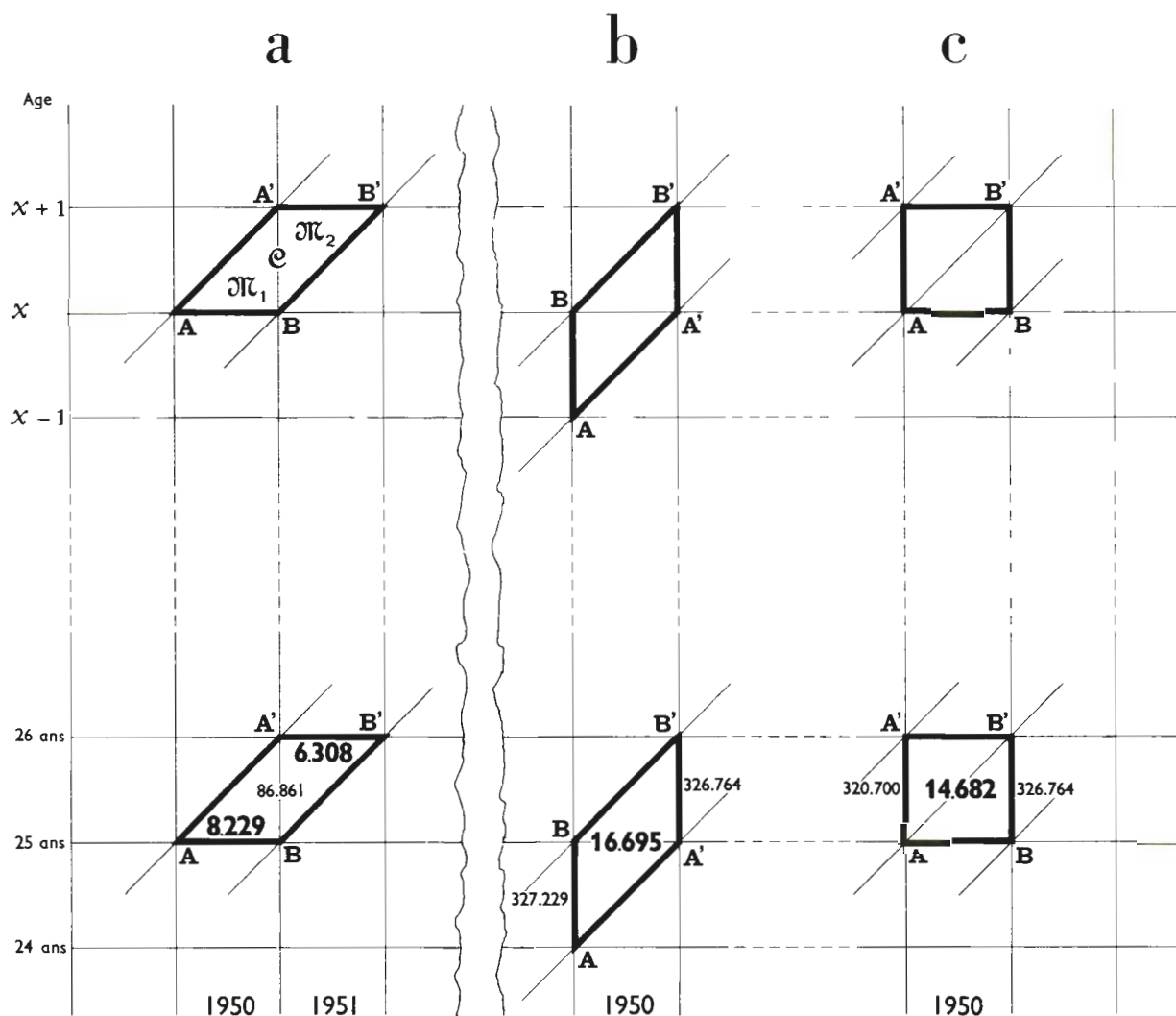


Figure 24



$$n_{25} = \frac{8.229 + 6.308}{86.861} = 167,4 \text{ ‰} \quad t'_{24} = \frac{16.695}{\frac{327.229 + 326.764}{2}} = 51,1 \text{ ‰} \quad t'_{25} = \frac{14.682}{\frac{320.700 + 326.764}{2}} = 45,4 \text{ ‰}$$

Figure 25 - Calculs sur la nuptialité des célibataires. Exemples empruntés à la population féminine française.

Par ailleurs, l'effectif total de la génération à cet anniversaire, soit $\mathfrak{S}_x$ vérifie

$$\mathfrak{S}_x = \mathfrak{S}_{15} \frac{S_x}{S_{15}}$$

Il en découle que la proportion des célibataires à l'anniversaire x , soit

$$\frac{c_x}{\mathfrak{S}_x} \quad (P)$$

s'exprime en fonction de données des tables de mortalité (des célibataires et de l'ensemble de la génération), et de nuptialité selon la formule,

$$\frac{c_x}{\mathfrak{S}_x} = \frac{C_x}{C_{15}} \cdot \frac{S_k}{S_x}$$

Par conséquent, la quantité $\frac{C_x}{C_{15}}$, qui fournit la série des célibataires de la table, vaut

$$\frac{C_x}{C_{15}} = \frac{c_x}{S_x} \frac{S_x}{S'_x} \quad (E)$$

Ainsi, la proportion (P) s'identifie aux célibataires de la table de nuptialité au facteur correctif

$$\frac{S_x}{S'_x}$$

près. En raison de la surmortalité de célibataires, ce facteur est légèrement supérieur à 1 : dans la mesure où nous négligeons le facteur correctif, nous surestimons la nuptialité.

Les résultats que nous venons d'établir, l'ont été aux anniversaires x ; ils s'étendent évidemment aux groupes d'un même âge en années révolues ; nous avons réservé la notation x' pour désigner l'âge de ces groupes, en sorte que la forme (E) se transforme en

$$\frac{C_{x'}}{C_{15}} = \frac{c_{x'}}{S_{x'}} \frac{S_{x'}}{S'_{x'}} \quad (E')$$

Cette formule s'étend encore à des groupes de plusieurs générations, par exemple des groupes quinquennaux ; si l'on dispose par ailleurs d'une suite de recensements quinquennaux, on peut suivre ainsi commodément des groupes de cinq générations et décrire leur nuptialité (nuptialité des célibataires).

A titre d'exemple, nous amorçons dans le tableau 27, l'histoire de la génération féminine française 1930 en utilisant les proportions de célibataires. La dernière colonne du tableau donne des estimations des quantités C_x , probablement un peu surévaluées pour la raison qui a été indiquée ; toutefois, les effets de cette surévaluation sont d'autant plus sensibles que l'âge est plus élevé, (puisque ce sont les quantités S_x qui interviennent) ; ils sont peu importants aux jeunes âges, époque où la nuptialité est très intense.

Tableau 27

Proportions de célibataires dans la génération féminine française 1930

Date	Age (en années révolues)	Effectif total de la génération $\mathfrak{S}$	Célibataires c	$\frac{c}{\mathfrak{S}}$ (p. 10.000)
1. 1. 1946	15 ans	331.177	329.569	9.951
1. 1. 1947	16 ans	330.714	327.112	9.891
1. 1. 1948	17 ans	330.327	319.095	9.660
1. 1. 1949	18 ans	330.304	301.518	9.128
1. 1. 1950	19 ans	330.311	272.041	8.236
1. 1. 1951	20 ans	329.936	234.694	7.113
1. 1. 1952	21 ans	329.587	195.469	5.931
.....				

Le tableau 28 offre l'exemple des résultats auxquels on peut parvenir en rapprochant les données de recensements quinquennaux successifs. L'exemple a été choisi à dessein pour souligner un des écueils fréquents de cette méthode : la qualité insuffisante des données qu'elle utilise, notamment la déclaration d'état matrimonial ; de fait, dans le tableau 28, le dernier chiffre de la dernière colonne, par exemple, est très suspect (le comparer au chiffre précédent)⁽¹⁾ : naturellement ces insuffisances ont quelque chance d'être particulièrement accusées du fait qu'il s'agit de recensements anciens.

(1) Le brusque passage de 35.118 à 19.506 dans la colonne des $\mathfrak{S}$, puis la remontée à 30-40 ans sont aussi curieux...

Tableau 28

Proportion de célibataires dans le groupe de générations féminines 1851-1855
du département des Côtes du Nord

Date*	Age	Effectif total du groupe de géné- rations §	Célibataires c	$\frac{c}{§}$ (p. 10.000)
.....				
1876	20-24 ans	35.118	25.988	7.400
1881	25-29 ans	19.506	9.045	4.637
1886	30-34 ans	19.876	6.854	3.448
1891	35-39 ans	19.637	5.001	2.547
1896	40-44 ans	17.755	4.178	2.353
1901	45-49 ans	16.068	3.005	1.870
.....				
* Date du recensement				

En principe, le travail d'élaboration ne s'arrête pas aux calculs de la dernière colonne des tableaux 27 et 28 ;

- avec les chiffres du tableau 27 on peut se proposer d'obtenir, par interpolation, les célibataires de la table classique (passage de $C_{x'}$ à C_x).

- il est beaucoup plus hasardeux d'entreprendre une telle estimation avec les résultats du tableau 28 ; on peut du moins résumer de façon satisfaisante la table de nuptialité que l'on ne saurait établir avec précision et pour cela on estimera les quantités C_{50} et m (âge moyen) par les formules

$$C_{50} = \frac{1}{2} \left(\frac{C_{45'-49'}}{§_{45'-49'}} + \frac{C_{50'-54'}}{§_{50'-54'}} \right)$$

$$m = \left[15 + 5 \left(\frac{C_{15'-19'}}{§_{15'-19'}} + \dots + \frac{C_{45'-49'}}{§_{45'-49'}} \right) - 50 C_{50} \right] \frac{1}{1 - C_{50}}$$

Tableau 29

Proportions des personnes inactives dans une génération masculine et description
de l'entrée en activité (table d'activité)

Age (en années révolues)	Proportion des personnes inactives (p. 1.000) $I_{x'}$	Nouvelles entrées en activité $a(x', x' + 1)$	quotients $\frac{a(x', x' + 1)}{I_{x'}}$
13 ans	1.000	250	0,25
14 ans	750	182	0,24
15 ans	568	140	0,25
16 ans	428	112	0,26
17 ans	316	96	0,30
18 ans	220	70	0,32
19 ans	150	50	0,33
20 ans	100	...	
.....	...	...	

Les proportions aux divers âges ou durées, des personnes qui dans une génération ou une cohorte quelconque n'ont pas encore vécu un événement d'un type donné (pas nécessairement démographique), constituent toujours des indices intéressants, susceptibles d'être utilisés ainsi que nous venons de le faire. Cette remarque rejoint d'ailleurs celle, plus générale, par laquelle nous avons conclu le chapitre II (cf. page 59), et qui signalait le large domaine d'application des modèles descriptifs présentés tout au long de ce chapitre. Les proportions en question sont précisément un des moyens les plus commodes pour aboutir à ces types de description.

Nous donnons un exemple dans le tableau 29 : il s'agit de l'entrée en activité ; les proportions de personnes inactives (compléments à 1 des taux d'activité communément employés), ont encore la signification des survivants⁽¹⁾. A partir de là, nous avons ajouté les deux colonnes (événements et quotients) qui constituent avec des I_x , la table d'activité ; c'est encore la série $\{a(x', x' + 1)\}$ des événements qui est la plus intéressante.

Un exemple synthétique

Arrivé à ce point de notre exposé, il est bon de synthétiser les trois principales méthodes d'établissement d'histoires de cohortes, à partir des statistiques démographiques courantes, que nous verrons de présenter.

Nous utiliserons pour cela l'ensemble des données d'une observation continue parfaite dont nous détacherons tel ou tel type de données, correspondant aux différents genres d'informations que l'on peut posséder dans la pratique ; nous travaillerons sur des résultats numériques auxquels nous avons déjà eu recours en partie (tableau 23)⁽²⁾, ceux correspondant aux remariages d'une catégorie de veufs ; voici des données (tableau 30), limitées aux cinq premières années.

Tableau 30
Données sur une cohorte de veufs

Durée écoulée depuis le veuvage x	Survivants dans la cohorte $\mathfrak{S}_x$	Remariages $\mathcal{R}(x, x + 1)$	Décès	
			de veufs $\mathcal{D}(x, x + 1)$	de remariés $\mathcal{D}'(x, x + 1)$
0 an	500	150	10	0
1 an	490	80	8	2
2 ans	480	40	8	2
3 ans	470	20	6	4
4 ans	460	10	6	4
5 ans	450			
.....	...			

Trois possibilités sont donc à envisager

a) on possède les données permettant de calculer des quotients (ce qui correspond aux cas de figure 23-a et 25-a) ; avec les n' auxquels on aboutit (cf. calculs du tableau 23), on parvient, par un calcul classique, aux $r(x, x + 1)$ du tableau 31 (1er calcul).

b) on ne peut calculer que des taux sur l'ensemble de la cohorte (taux de 2ème catégorie) ; on en déduit les $r'(x, x + 1)$ (2ème calcul du tableau 31) par la formule

$$r'(x, x + 1) = \frac{\mathcal{R}(x, x + 1)}{\frac{\mathfrak{S}_x + \mathfrak{S}_{x+1}}{2}}$$

(1) Nous avons d'ailleurs employé d'emblée une notation (I_x) qui évoque cette interprétation, alors qu'il s'agit rigoureusement de proportions nécessitant éventuellement une correction pour représenter la série cherchée ; signalons encore que le retour à l'inactivité peut fausser la description ou du moins lui faire prendre un sens différent (ce phénomène est surtout sensible chez les jeunes femmes qui quittent leur emploi en proportion appréciable après le mariage et les maternités).

(2) Données extraites de Roland Pressat. *Pratique de la démographie. Trente sujets d'analyse*. Sujet n° 16 ; le texte de ce sujet est lui-même bâti sur des données de Louis Henry.

c) enfin on ne connaît que le partage en personnes remariées et personnes encore veuves, aux différents anniversaires du remariage ; on en déduit les $r''(x, x + 1)$ (3ème calcul du tableau 31) par l'intermédiaire des proportions

$$\frac{v_x}{\mathfrak{S}_x}$$

Tableau 31

Divers calculs de remariages de veufs en l'absence de mortalité

Anniversaire du veuvage x	1er calcul			2ème calcul			3ème calcul			
	V_x	n'_x (p. 1000)	$r(x, x + 1)$	$\frac{\mathfrak{S}_x + \mathfrak{S}_{x+1}}{2}$	$\mathcal{R}(x, x + 1)$	$r'(x, x + 1)$	$\mathfrak{S}_x$	v_x	$\frac{v_x}{\mathfrak{S}_x}$ (p. 1000)	$r''(x, x + 1)$
0	1000	303	<u>303</u>	495	150	<u>303</u>	500	500	1000	<u>306</u>
1	697	238	<u>166</u>	485	80	<u>165</u>	490	340	694	<u>169</u>
2	531	161	<u>85</u>	475	40	<u>84</u>	480	252	525	<u>91</u>
3	446	100	<u>45</u>	465	20	<u>43</u>	470	204	434	<u>47</u>
4	401	57	<u>23</u>	455	10	<u>22</u>	460	178	387	<u>27</u>
5	378		<u>622</u>			<u>617</u>	450	162	360	<u>640</u>

Seul, le premier calcul est rigoureux. Ainsi que nous l'avons vu, le 2ème calcul sous-estime l'importance du phénomène tandis que le 3ème le surestime, cela dans le cas où la mortalité des personnes n'ayant pas encore subi le phénomène est supérieure à celle des personnes l'ayant subi. C'est le cas avec les données du tableau 30 et les résultats du tableau 31 vérifient bien ce que les développements théoriques ont annoncé : en totalisant les remariages on trouve le classement, par intensité de nuptialité croissante,

2ème calcul, 1er calcul, 3ème calcul

La fécondité

La description de la fécondité sans distinction de rang de naissance requiert la détermination des taux u_x (description dans une promotion de mariages), et f_x (description dans une génération).

Le calcul de f_x est possible dans la majorité des cas : on se trouve dans un des trois cas de la figure 22 et aucun problème particulier ne se pose.

Par contre, la détermination de u_x n'est pour ainsi dire jamais possible car on ne possède pas de statistiques de mariages subsistants selon la durée du mariage ; on est donc contraint d'utiliser la seule donnée chiffrée que l'on possède, l'effectif initial de ces mariages. Autrement dit, on renonce à corriger le phénomène étudié (la fécondité des mariages) des phénomènes perturbateurs (désunions et éventuellement migrations).

Le tableau 32 offre l'exemple d'un tel calcul.

Pour essayer de mieux approcher le phénomène à l'état pur, on peut estimer ces effectifs de mariages subsistants, que les statistiques ne donnent pas, en faisant intervenir des lois convenables de mortalité des conjoints et de divorcialité. Mais, lorsque l'on procède à des comparaisons dans un même pays à des dates rapprochées, une telle correction n'est pas indispensable, l'influence des phénomènes perturbateurs étant à peu près la même aux différentes dates considérées. L'écart entre une mesure sans élimination des perturbations et une mesure du phénomène pur est d'ailleurs faible (quelques naissances pour 100 mariages), car le phénomène se manifeste, pour l'essentiel, dans des délais assez courts après le mariage.

Tableau 31/2

Descendance de la promotion de mariages française 1946 sans élimination des événements perturbateurs (désunions et migrations) (effectif initial : 516.882 mariages)

Année des naissances	Naissances		Descen- dance pour 1000 mariages	Année des naissances	Naissances		Descen- dance pour 1000 mariages
	Nombre absolu	Pour 1000 mariages			Nombre absolu	Pour 1000 mariages	
1946	58.840	114	114	1955	44.607	86	1.844
1947	220.261	426	540	1956	39.071	76	1.920
1948	139.556	270	810	1957	33.242	64	1.984
1949	117.462	227	1.037	1958	27.827	54	2.038
1950	102.459	198	1.235	1959	24.290	47	2.085
1951	84.112	163	1.398	1960	19.836	38	2.123
1952	71.895	139	1.537	1961	16.937	33	2.156
1953	60.968	118	1.655	1962	13.464	26	2.182
1954	53.161	103	1.758	1963	11.108	21	2.203