CONCURRENCE ET INDÉPENDANCE ENTRE PHÉNOMÈNES DÉMOGRAPHIQUES
Réflexions et commentaires sur un article de Xavier Thierry*

Les enquêtes démographiques sont depuis longtemps « biographiques », mais elles se contentaient initialement des chronologies génésique et contraceptive des femmes interrogées. Seulement dans les deux dernières décennies se sont multipliés et systématisés les recueils de biographies croisées associant des dimensions variées de la vie individuelle : cursus scolaire, carrière professionnelle, parcours résidentiel, histoire familiale, etc. L'intrication de ces calendriers a nécessité le développement d'outils d'analyse spécifiques, dont Daniel Courgeau et Éva Lelièvre ont illustré le maniement dans leurs exploitations de l'enquête « 3 B - Triple biographie » et dont ils ont exposé les principes dans un manuel.

L'évolution des comportements conjugaux crée aujourd'hui de nouvelles difficultés à l'analyse démographique : la diffusion de formes alternatives de vie en couple venant concurrencer le mariage. La présentation de ces difficultés et d'éléments pour les surmonter, par Xavier Thierry dans un article récent, a provoqué une réaction de Daniel COURGEAU** *et* Éva LELIÈVRE** *qui montrent le parallélisme des problèmes posés par la concurrence entre les formes d'union et par l'interaction entre les phénomènes démographiques, analysée dans leur manuel. Le parallélisme des problèmes entraîne évidemment celui des solutions. Les conclusions diffèrent de celles de Xavier Thierry. Il nous semble que s'ouvre là une discussion méthodologique intéressante ; nous avons joint nos réflexions après celles de Daniel Courgeau et Éva Lelièvre et souhaitons que d'autres lecteurs s'associent à ce débat.*

---

* Xavier Thierry, « La nuptialité à l'épreuve de la cohabitation », *Population*, 4, 1993, 939-974.

** INED

RISQUES COMPÉTITIFS ET INDÉPENDANCE : CADRE THÉORIQUE D'UNE RÉFLEXION

Daniel COURGEAU, Éva LELIÈVRE

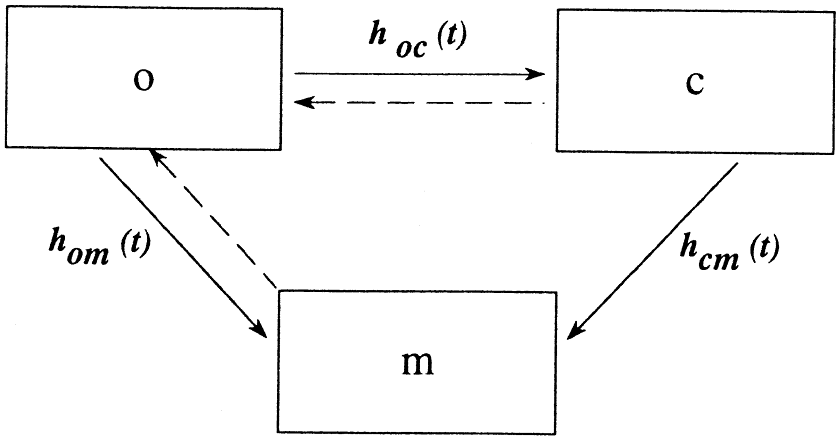
L'étude de la nuptialité ne suffit plus pour rendre compte des modalités de la mise en couple, du fait de la diffusion massive de la cohabitation. La première mise en union peut en effet prendre la forme de deux événements mutuellement exclusifs. La concurrence entre ceux-ci doit être modélisée. C'est ce que réalise l'approche longitudinale biographique. Celle-ci a l'avantage de proposer un cadre théorique solide dans le respect des principes de l'analyse démographique [Keilman, 1993]. Pressat le soulignait déjà en 1966 « la recherche des corrélations entre phénomènes démographiques si elle constitue encore un domaine inexploré, devrait pouvoir enrichir considérablement notre connaissance ».

La mise en union au sein d'une population de célibataires[1] se fait selon deux modalités : mariage direct ou cohabitation (celle-ci suivie éventuellement d'un mariage). Ces deux événements sont concurrents et les individus sont soumis au risque de ces deux modalités de mise en union. En démographie, en particulier si l'on se réfère à l'étude de la mortalité par cause, ce n'est pas une situation nouvelle du point de vue méthodologique [Aalen, 1978, Courgeau et Lelièvre, 1989]. Cependant dans le cas de la mise en union, la cohabitation peut être le prélude d'un mariage à venir (période de cohabitation pré-maritale). Il s'introduit donc une relation d'asymétrie : la cohabitation laisse ouverte la possiblité d'un mariage entre les partenaires, que le mariage direct soustrait ceux-ci au risque d'une cohabitation informelle.

L'article de X. Thierry qui traite de cette situation ne postule pas cette asymétrie. L'auteur explore en effet la validité d'une hypothèse d'indépendance entre les deux phénomènes, hypothèse que le constat précédent invalide sans nécessiter de plus ample « démonstration ». Cependant, si l'exercice est à notre sens sans objet, la question de l'interaction entre phénomènes, du traitement des phénomènes compétitifs et des divers types de dépendance est, elle, fondamentale : d'où cette note.

Un schéma simple Il est possible de représenter les divers états par trois cases (hors union (*o*), cohabitant (*c*) et marié (*m*) et les passages d'un état vers l'autre par des flèches, lorsque ces passages sont possibles. Nous avons également indiqué les quotients instan-

[1] On se limite ici, pour la clarté de l'exposé, à la première union de célibataires étant entendu que le raisonnement est transposable aux unions successives moyennant quelques adaptations. Comme dans l'article de Xavier Thierry, on se place dans le cas d'une analyse de données d'enquête rétrospective.



Schéma

tanés de passage : nous les définirons plus précisément tout au long de cette note.

L'asymétrie signalée précédemment se traduit ici par le fait que le passage de marié à cohabitant est impossible pour les conjoints que leur première mise en couple a unis.

L'analyse peut avoir plusieurs objectifs :

1) Mesurer et comparer l'intensité et le calendrier de chacun des types de mise en union des célibataires.

2) Mesurer et comparer la nuptialité des célibataires et des cohabitants, des célibataires selon qu'ils cohabitent ou non, c'est-à-dire mesurer l'effet de la cohabitation sur le mariage.

3) Évaluer l'effet qu'a la distribution marginale d'un des types de mise en union sur la mise en couple en général.

Comme on le verra, aucun de ces objectifs ne nécessite de faire l'hypothèse d'indépendance entre cohabitation et mariage.

Effets perturbateurs et indépendance

Pour mesurer des phénomènes démographiques à « l'état pur » [Henry, 1970], on doit éliminer l'effet des phénomènes perturbateurs (mortalité et migration dans l'étude de la nuptialité, par exemple) en faisant l'hypothèse d'indépendance avec le phénomène étudié. Même si cette hypothèse n'est pas vérifiée, les résultats n'en sont pas trop affectés tant que les événements perturbateurs sont en petit nombre par rapport à l'événement étudié. En effet, l'incidence de la mortalité et de la migration sur la nuptialité est généralement très faible si l'on conduit une étude à l'échelle nationale [Thierry, p. 947].

Néanmoins lorsqu'il s'agit d'événements en compétition les uns avec les autres, la situation devient très différente. La confusion provient de cette spécificité des événements compétitifs vis-à-vis des événements en interaction. Ainsi rappelons rapidement le cadre d'ensemble de l'étude de deux phénomènes démographiques.

Le schéma d'analyse que nous présentons dans la première partie de cette note, correspond au cas plus général de *l'étude des phénomènes (ou risques) compétitifs*, dont l'exemple classique est l'étude de la mortalité par causes : de même qu'un individu peut soit se marier, soit cohabiter; un survivant peut mourir d'un accident de la route, d'une maladie cérébro-vasculaire, etc. Dans tous les cas l'individu est soumis aux divers risques, mais ne peut en connaître qu'un seul.

Ce schéma est complètement différent de celui de l'étude d'un même phénomène selon l'état dans lequel se trouve l'individu au moment où il le connaît : mortalité des actifs et des retraités, par exemple. Ce second schéma est celui de *l'étude d'un même phénomène avec des caractéristiques dépendant du temps*. C'est ce schéma que nous présenterons dans la dernière partie de cette note : l'individu est soumis ici seulement au risque de mariage, selon qu'il est lui-même déjà cohabitant ou toujours hors union. Dans tous ces cas l'état final est le même (marié), mais c'est l'état initial qui se diversifie.

Enfin un troisième schéma apparaît qui correspond à *l'étude de deux ou plusieurs phénomènes en interaction* : nuptialité et départ de l'agriculture par exemple. Dans ce cas aucun des phénomènes n'empêche l'apparition des autres contrairement à ce qui se passe dans l'étude de la cohabitation et du mariage, où le mariage empêche que toute cohabitation puisse le suivre directement. Nous avons déjà présenté ce type d'analyse de façon détaillée et renvoyons le lecteur aux chapitres rédigés sur ce schéma (Courgeau et Lelièvre, 1989, chapitre V et VI).

Si l'on veut analyser la cohabitation et le mariage dans nos sociétés actuelles, la formulation classique de l'analyse de la nuptialité va s'en trouver modifiée. En effet, un des objectifs de l'analyse va justement consister à évaluer les effets conjoints des deux phénomènes, les formulations proposées par l'analyse probabiliste évitant de faire l'hypothèse coûteuse d'indépendance[(2)].

Définition et estimation des quotients

Poursuivant notre premier objectif, considérons d'abord le couple de variables aléatoires (T, U) où T représente la durée jusqu'à la mise en union et U le type d'union : égal à m si l'individu se marie, égal à c s'il devient cohabitant. Dans ce cas, on définit les quotients instantanés suivants :

$$h_{om}(t) = \lim_{dt \to 0} \frac{P(T < t + dt, U = m \mid t \leq T)}{dt} \qquad [1]$$

[(2)] Du point de vue sociologique, l'interprétation d'indépendance de la cohabitation et du mariage est totalement absurde.

et

$$h_{oc}(t) = \lim_{dt \to 0} \frac{P(T < t + dt, U = c \mid t \leq T)}{dt} \qquad [2]$$

Comme le soulignait R. Kasparian (1993) «du fait qu'il est instantané, ce quotient contrairement au quotient classique annuel ou quinquennal n'est pas affecté par l'interférence d'autres événements». Ces quotients, qui correspondent à chaque événement en présence de l'autre, sont définis sans aucune hypothèse d'indépendance entre événements. Il y a diverses possibilités de les estimer en fonction des données dont on dispose et justement des hypothèses que l'on fait.

Si l'on ne dispose que du nombre de mariages et de mises en cohabitation au cours de l'année, on peut faire l'hypothèse qu'au cours de cette année a, les événements se produisent de façon uniforme dans le temps. On montre que l'estimation du maximum de vraisemblance du quotient de nuptialité est dans ce cas :

$$\hat{h}_{om}(a) = \frac{n_m}{n_m + n_c} \log\left[1 - \frac{n_m + n_c}{r_o}\right] \qquad [3]$$

où n_c est le nombre d'entrées en cohabitation enregistrées l'année a, n_m le nombre de mariages et r_o la population en début d'année a des individus dans l'état «hors union». Si le rapport $(n_m + n_c) / r_o$ est faible on peut montrer que l'estimation s'écrit plus simplement :

$$\hat{h}_{om}(a) = \frac{n_m}{r_o - \frac{1}{2}(n_m + n_c)} \qquad [3bis]$$

avec une estimation semblable pour $\hat{h}_{oc}(a)$. On a bien entendu la possibilité d'estimer les variances de ces quotients.

La connaissance de ces quotients instantanés permet de calculer la probabilité qu'un des événements se produise, se marier au cours de l'année a, par exemple. Elle est égale à :

$$p_{om}(a) = 1 - \exp(-h_{om}(a)) \text{ estimée par } \hat{p}_{om}(a) = \frac{n_m}{r_o - \frac{1}{2} n_o} = QC_O(a) \qquad [4]$$

Cette estimation classique, rappelée par X. Thierry (tableau 2, p. 949) n'est pas un quotient instantané mais une probabilité. Le dénominateur prend en compte la population en début d'intervalle moins la moitié des entrées en cohabitation ; sa détermination conduit à supposer cette fois que les deux événements, mariage et cohabitation, sont indépendants.

Supposons maintenant que l'enquête rétrospective soit à même de fournir les dates précises de mise en union au cours de l'année a. Une des principales innovations apportées par Aalen (1978) en généralisant le cadre théorique proposé par Kaplan et Meier (1958), est d'offrir la possibilité d'estimer l'occurrence d'une échéance simple (la mise en union) en envisageant plusieurs modalités de cette échéance (cohabitation, mariage) sans faire d'hypothèse de constance au cours de l'année des différents risques ni *a fortiori* d'indépendance entre les 2 modalités de mise en union.

Dans ce cadre l'estimateur du quotient instantané s'écrit :

$$\hat{h}_{oi}(t) = \frac{n_i(t)}{Y_o(t)} \quad \text{estimateur de Aalen à l'instant } t \text{ dans l'année } a \qquad [5]$$

$$\hat{h}_{oi}(a) = \sum_{t \in [a, a+1[} \frac{n_i(t)}{Y_o(t)} \quad \text{estimateur de l'intensité cumulée sur l'année } a \qquad [6]$$

où $n_i(t)$ est le nombre d'événements de type i, et $Y_o(t)$ la population soumise au risque dans l'état « hors union » juste avant l'instant t. On dispose ici encore d'une estimation simple des variances de ces quotients [Courgeau et Lelièvre, 1989].

La théorie montre, de plus, que le passage du temps continu (on dispose pour chaque individu des durées de séjour exactes au cours de l'année a, ce qui est le cas dans une enquête rétrospective) au temps discret (calcul de quotients par âge) conduit à de résultats peu différents si l'on prend garde de maintenir la durée dans les limites habituelles(3) : une année environ [Argas et Kangas, 1992]. Le calcul de quotients quinquennaux moyens à partir de quotients instantanés estimés sur une période d'un an reste par contre complexe du fait des interférences avec l'événement perturbateur [Kasparian, 1993].

D'autre part, la loi d'additivité des probabilités donne le quotient marginal estimant l'incidence de la mise en union au sein d'une population soumise au risque :

$$h(t) = h_{om}(t) + h_{oc}(t) \qquad [7]$$

et de façon semblable la densité marginale et les fonctions de séjour marginales [Cox et Oakes, 1984]. Ainsi sachant qu'un événement s'est produit à l'instant t, la probabilité pour que celui-ci soit, par exemple, un mariage, est donnée par $h_{om}(t) / h(t)$. Cette probabilité étant celle que l'on cherche à évaluer si l'on poursuit l'objectif 3. De sorte que la probabilité marginale de connaître un mariage est :

$$P(U=m) = \int_{t=0}^{t_f} h_{om}(t) \exp\left[-\int_{\theta=0}^{t} h(\theta)\, d\theta \right] dt \qquad [8]$$

où t_f est l'âge maximum auquel les individus sont observés dans l'enquête. On calcule de façon semblable la probabilité marginale de connaître une cohabitation comme première union.

---

(3) Voir Courgeau et Lelièvre (1989) p. 28 le sous-chapitre « Temps discret, temps continu ».

Synthèse des quotients Une fois les quotients estimés, il est naturel de chercher à avoir une information synthétique sur le phénomène étudié. Ainsi, on peut calculer des proportions de premières unions débutées en cohabitation ou en mariage [Thierry, p. 152], ce qui correspond à calculer des fonctions de séjour. Mais cela n'a de signification cette fois encore, que si les risques sont indépendants. Ce qui dans le cas qui nous occupe n'est sans doute pas vérifié.

Si l'on utilise la théorie probabiliste dans les cas de risques compétitifs, les quotients cumulés sont en revanche intéressants à considérer. Ils se définissent comme une intégrale :

$$H_{oi}(t) = \int_{\theta=0}^{t} h_{oi}(\theta)\, d\theta \quad \text{estimée à l'âge } a \text{ par } \hat{H}_{oi}(a) = \sum_{x=0}^{a} up\, \hat{h}_{oi}(x) \qquad [9]$$

L'avantage de calculer ces quotients cumulés ($\hat{H}_{om}(t)$, $\hat{H}_{oc}(t)$) est que, contrairement aux fonctions de séjour, ils peuvent être considérés comme des processus indépendants, même lorsque les risques sont dépendants [Aalen, 1978 ; Andersen *et al.*, 1992]. Il en résulte, par exemple, que les représentations graphiques de ces quotients cumulés en fonction du temps (graphiques de Nelson-Aalen) pour chaque risque peuvent être considérés indépendamment les uns des autres, sans confusion possible. Ces quotients cumulés permettent donc de mesurer et comparer l'intensité et le calendrier de chacun des types de mise en union.

Nous avons aussi répondu à l'objectif 1, toutes ces fonctions pouvant être estimées sans aucune hypothèse sous-jacente, les risques étant considérés conjointement. On peut ensuite essayer de voir l'effet de diverses caractéristiques individuelles (niveau d'éducation, origines sociales, profession, etc.) sur ces quotients en utilisant par exemple un modèle à risques proportionnels [Hoem, 1986].

Est-il possible d'analyser certains risques en l'absence des autres ? Lorsque l'on cherche à aller plus loin, en essayant de mesurer l'intensité du premier mariage ou de la cohabitation en l'absence de tout phénomène perturbateur (« comme s'il n'existait que cette seule forme d'union possible » [Thierry, p. 948]), il convient d'opérer avec prudence. Voyons d'abord comment formaliser cette approche.

Supposons que, pour chaque individu, il existe deux variables aléatoires : T_m qui représente l'instant du mariage en l'absence de cohabitation et T^c qui représente l'instant d'entrée en cohabitation en l'absence de mariage. En fait, on n'observe pas directement l'occurrence de ces variables aléatoires, mais celles de $T = \min(T^m, T^c)t$ de U, égale à m si $T = T^m$ et égale à c si $T = T^c$.

Avec ces nouvelles notations les quotients précédents peuvent s'écrire de façon différente :

$$h_{om} = \lim_{dt \to 0} \frac{P\,(T^m < t + dt \mid t \leq T^m, t \leq T^c)}{dt} \qquad [10]$$

avec une formulation semblable pour $h_{oc}\,(t)$. En revanche, le quotient relatif à la variable aléatoire T^m, correspondant au mariage en l'absence de cohabitation, s'écrit :

$$h^{om}\,(t) = \lim_{dt \to 0} \frac{P\,(T^m < t + dt \mid t \leq T^m)}{dt} \qquad [11]$$

avec une formulation semblable pour h^{oc} (t).

On voit facilement que les probabilités qui interviennent dans les deux types de quotients sont conditionnées par des événements différents : dans le premier cas [10], l'individu est hors union, tandis que dans le second cas [11] il est seulement non marié. Par suite, ces quotients ne peuvent être égaux que si les variables T^m et T^c sont indépendantes[4].

Tester l'interaction entre deux phénomènes

Dans le cas de l'étude de la cohabitation et du mariage et poursuivant l'objectif 2, il est possible d'analyser plus avant la dépendance entre ces deux phénomènes, mais dans un sens seulement : de la cohabitation vers le mariage puisque certains cohabitants finissent par s'épouser. Pour ce faire, il nous faut considérer une troisième série de quotients instantanés.

$$h_{cm}\,(t,l) = \lim_{dt \to 0} \frac{P\,(T^m < t + dt \mid t \leq T^m, T^c = l)}{dt} \quad \text{avec } l \leq t \qquad [12]$$

Ce quotient dépend, non seulement de la date t du mariage (l'âge de l'individu à ce moment), mais également de celle à l'entrée en cohabitation, l. Si les observations sont suffisamment nombreuses, ces quotients sont à nouveau faciles à estimer à partir des données d'une enquête rétrospective. Ils permettent de vérifier les liens entre cohabitation et mariage.

Supposons, par exemple, que ce quotient ne dépende que de l'âge de l'individu. On peut alors tester l'égalité suivante :

$$h_{om}\,(t) = h_{cm}\,(t) \qquad [13]$$

Si cette égalité est vérifiée quel que soit t, on peut seulement affirmer qu'il y a indépendance unilatérale [Courgeau et Lelièvre, 1989], entre cohabitation et mariage : ce qui reviendrait à dire qu'une cohabitation antérieure ne modifie en rien les probabilités de mariage de l'individu. On sait d'autre part, que le mariage direct empêche une première cohabitation. Il

---

(4) En fait, une condition un peu moins stricte est nécessaire : la quasi-indépendance [Cox et Oakes, 1984]. Cela revient à dire que les probabilités conditionnées par $T \leq T^c$ et $T \leq T^m$ sont égales à celles conditionnées par $T \leq T^m$ seulement.

est dès lors impossible de déceler l'indépendance totale entre T^m et T^c, car celle-ci doit être vérifiée dans les deux sens.

En revanche, la comparaison de $h_{om}(t)$ et de $h_{cm}(t,l)$, le mariage des célibataires et des cohabitants en tenant compte de la durée de cohabitation doit être poursuivie [Andersen *et al.*, 1992]. On peut également comparer $h_{oc}(l) \cdot h_{cm}(t,l)$ et $h_{om}(t)$: la nuptialité des célibataires selon qu'ils cohabitent ou non. Elles sont du plus grand intérêt pour l'analyse des relations entre cohabitation et mariage. Cela permet de répondre à l'objectif 2.

Conclusion Tant il est possible d'éliminer des phénomènes perturbateurs tels que la mortalité ou les migrations internationales dont les interférences sont finalement très faibles vis-à-vis du phénomène étudié, tant il paraît sans objet et illégitime de faire l'hypothèse d'indépendance dans les cas bien identifiés de phénomènes compétitifs.

Les modèles à risques compétitifs, rappelés ici, permettent dans le strict respect des principes de l'analyse démographique, de prendre en compte la non-dépendance de tels processus. Dans le cas de l'étude du mariage et de la cohabitation, l'objectif est bien de tester les interactions existant entre ces deux phénomènes ainsi que l'influence de la cohabitation sur l'intensité de la mise en union. Par l'intermédiaire des quotients instantanés et des quotients cumulés pour lesquels on peut calculer les variances (ce qui est nécessaire dans l'utilisation de données d'enquêtes), il est possible de répondre clairement à ces objectifs.

Daniel COURGEAU, Éva LELIÈVRE

RÉFÉRENCES

AALEN, O., 1978 – « Nonparametric inference in connection with multiple decrement models », *Scandinavian Journal of Statistics*, 3, 15-27.

ANDERSEN, P.C., BORGAN, O. GILL, R.D., KEIDING, N., 1993 – *Statistical models based on counting processes*. Springer-Verlag, New York, 768 p.

ARJAS, E., KANGAS, P., 1992 – « A discrete time method for the analysis of event histories », in Trussel, Hankinson, Tilton, eds. *Demographic applications of event history analysis*. Oxford University Press, Oxford, 253-266.

COURGEAU, D., LELIEVRE, E, 1989 – *Analyse démographique des biographies*, Editions de l'INED, Paris, 268 p. (*Event history analysis in demography*, 1993, Oxford University Press, Oxford, 226 p.).

COX, D. R., OAKES, D., 1984 – *Analysis of survival data*. Chapman and Hall, London, 202 p.

HENRY, L. 1972 – *Démographie. Analyse et modèles*. Larousse, Paris, 342 p.

HOEM, J., 1986 – « The impact of education on modern union dissolution. » *European Journal of Population*, 2, 2, 113-137.

KAPLAN, E., MEIER, P., 1958 – « Nonparametric Estimation from Incomplete Observations », *JASA*, vol.53, 457-481.

KASPARIAN, R., 1993 – « L'analyse longitudinale de la population active : une typologie des profils de carrière des générations françaises de 1911 à 1965 ». *Population*, 48, 3, 627-654.

KEILMAN, N., 1993 – « Emerging issues in demographic methodology », in *European Population II Demographic dynamics*, Blum & Rallu eds. John Libbey, INED, Paris, 483-508.

THIERRY, X., 1993 – « La nuptialité à l'épreuve de la cohabitation. » *Population*, 48, 4, 939-974.

QUELQUES COMMENTAIRES SUR UN PROBLÈME FONDAMENTAL DE L'ANALYSE DÉMOGRAPHIQUE

Patrick Festy

La note de Daniel Courgeau et Éva Lelièvre est organisée autour de trois questions étroitement imbriquées, comme dans l'article de Xavier Thierry auquel elle se réfère :

— peut-on mesurer, à partir des observations sur les formes prises par la première union, une fréquence du mariage «en l'absence de mise en cohabitation» ?

— cette nuptialité directe se confond-elle avec celle résultant d'un mariage après cohabitation ?

— si les réponses aux deux premières questions sont positives, la nuptialité «directe» (mariage sans cohabitation préalable, mesuré en l'absence de cohabitation) sera égale à la «primo» nuptialité (mariage de célibataire) ; sinon qu'en sera-t-il ?

La possibilité de traiter les deux dernières questions est cependant soumise à la réponse donnée à la première. Aussi ce point initial mérite-t-il une attention privilégiée.

Mesure et combinaison des quotients annuels

Daniel Courgeau et Éva Lelièvre distinguent deux méthodes de mesure des quotients de nuptialité par mariage direct. La première se fonde sur des données agrégées par année ; elle a été utilisée par Xavier Thierry dans son article. Le quotient de nuptialité entre les anniversaires x et $x + 1$ s'obtient en rapportant aux célibataires hors union à l'anniversaire x les mariages directs observés entre x et $x + 1$ et ceux qu'ont empêchés des mises en union au même âge. L'hypothèse d'indépendance entre nuptialité et mise en cohabitation est introduite lorsqu'on estime les mariages empêchés en supposant que les nouveaux cohabitants se seraient mariés comme les autres personnes hors union.

La seconde méthode recourt à des données biographiques individuelles, où le calcul de l'âge exact à chaque événement est possible (en mois ou en jours, selon la précision des questionnaires). A chaque âge exact où des mariages directs se produisent, on peut dès lors calculer un quotient en rapportant ceux-ci au nombre de personnes encore hors union à la veille de cet âge. Dans le calcul de chaque quotient, on peut négliger l'effet perturbateur résultant des entrées en cohabitation concomitantes, tant celles-ci sont faibles par rapport au nombre de personnes hors union. A la suite de Allen, D. Courgeau et É. Lelièvre proposent de sommer ces quotients sur une année pour mesurer un quotient annuel, qui ne diffère guère de celui calculé par la première méthode.

On ne s'étonnera pas de ces résultats car les différences entre les deux méthodes sont très minces, tant que les événements étudiés sont peu nombreux à chaque âge par rapport au nombre de personnes hors union. On sait que le facteur correctif du quotient annuel classique a alors peu d'importance, *a fortiori* celui des quotients mensuels ou journaliers. En outre, la somme de ces derniers s'écarte à peine du résultat qu'on obtiendrait en passant par la multiplication de leur complément à l'unité.

D'éventuelles divergences entre les choix de X. Thierry et ceux de D. Courgeau et É. Lelièvre se trouvent donc ailleurs. Elles tiennent à la possibilité d'étendre au-delà d'un an les calculs et certaines approximations sans conséquences dans le cadre annuel. Chacun s'accorde en effet à reconnaître que la combinaison classique des quotients de nuptialité successifs, par produit de leur complément à l'unité, n'est possible que si on peut postuler l'indépendance entre le phénomène étudié (par exemple, la nuptialité par mariage direct) et le phénomène qui en empêche parfois l'expression (l'entrée en première union par cohabitation). Chacun admet en outre, *a priori* ou à l'issue d'un raisonnement par l'absurde, que cette condition n'est pas remplie dans la concurrence entre les formes alternatives de première union. Mais X. Thierry étudie alors la vraisemblance d'autres relations possibles entre les phénomènes concurrents et conclut à l'abandon de toute combinaison des quotients, alors que D. Courgeau et É. Lelièvre non seulement conservent le calcul des quotients, mais proposent leur cumul, comme mesure de l'intensité du phénomène étudié, et leur répartition par âge, comme mesure de son calendrier.

L'indépendance entre phénomènes : deux illustrations

Il n'est cependant guère possible d'argumenter dans un sens ou dans l'autre sans être revenu sur la notion d'indépendance, à l'origine de la divergence. Nous le ferons en donnant à ces rappels la forme de deux illustrations empiriques.

La première se fonde sur des données collectées par U. Cowgill et qui ont donné lieu à un commentaire par L. Henry[1]. Nous reprenons l'esprit de ce commentaire. U. Cowgill mesure la répartition des âges au décès dans des générations nées dans un village anglais au XVI^e^ siècle. L'âge moyen au décès est de l'ordre de 12 ans, mais il est calculé sur les seules personnes dont l'acte de décès a pu être retrouvé dans le même village, soit 36 % des naissances (tableau 1, colonne 1). Comme L. Henry, on peut supposer que les 64 % manquants sont des personnes décédées ailleurs, après avoir émigré. On peut imaginer la répartition des âges à la migration,

[1] Ursula M. Cowgill, « Life and death in the sixteenth century in the city of York », *Population Studies*, july 1967. Louis Henry, « Some comments on Ursula M. Cowgill's article », *Population Studies*, march 1968.

TABLEAU 1. – ÉLÉMENTS DE CONSTRUCTION D'UNE TABLE DE MORTALITÉ EN CAS DE FORTE ÉMIGRATION

Age	Décès observés (1)	Émigrants supposés (2)	Survivants présents* (3)	Quotients de mortalité (4)=(1)/(3)	Survivants* (5)	Quotients de mortalité $(6) = (1)/(3) - \frac{(2)}{2}$	Survivants* (7)
0-4	225	150	1 000	0,225	1 000	0,243	1 000
5-9	25	100	625	0,040	775	0,043	757
10-14	20	50	500	0,040	744	0,042	724
15-19	15	100	430	0,035	714	0,039	694
20-24	10	170	315	0,032	689	0,043	667
25-29	10	70	135	0,074	667	0,100	638
30-34	10	–	55	0,182	618	0,182	574
35+	45	–	45		506		470
Total	360	640					
Age moyen au décès**	12,2 ans				35,9 ans		34,2 ans

* au début de l'intervalle d'âge.
** l'âge moyen au décès est supposé égal à 1,25 an pour les 0-4 ans et 57,5 pour les 35 ans en plus.

en se fondant sur le profil habituel des pyramides de migrants, où prédominent les jeunes adultes et leurs jeunes enfants (tableau 1, colonne 2).

Négligeant le fait que les migrants partent tout au long d'un intervalle d'âge, et non à la fin de celui-ci, on peut mesurer les quotients quinquennaux de mortalité (tableau 1, colonne 4) en rapportant les décès observés (tableau 1, colonne 1) aux survivants encore présents au début de l'intervalle (tableau 1, colonne 3). Les survivants présents voient leur nombre très vite réduit, à mesure qu'on avance en âge, par les décès et surtout par les nombreuses émigrations.

La construction de la table de mortalité consiste à reconstituer, intervalle d'âge par intervalle d'âge, la disparition progressive d'une génération sous le seul effet des quotients de mortalité. Aux survivants d'un anniversaire, on applique la probabilité quinquennale de survie (complément à 1 du quotient quinquennal de mortalité), pour obtenir les survivants cinq anniversaires plus tard (tableau 1, colonne 5). Pour un même effectif à la naissance, les survivants des divers âges sont sensiblement plus nombreux que les survivants présents (tableau 1, colonnes 3 et 5), car ils réintègrent dans la population les émigrants qui seraient encore en vie sur leur lieu de naissance, s'ils n'avaient pas quitté le village. En appliquant à ce surcroît de population la probabilité de survie qu'on a calculée sur les seuls survivants présents, on postule que *ces émigrants auraient été exposés au même risque de mortalité que les autochtones, s'ils étaient restés sur place.* C'est l'hypothèse d'indépendance entre mortalité et migration.

Dans la table de mortalité ainsi calculée, l'espérance de vie à la naissance dépasse 35 ans. Une hypothèse plus vraisemblable sur le temps de présence des émigrants dans l'intervalle d'âge où ils quittent le village ne modifie ce résultat que faiblement (tableau 1, colonnes 6 et 7).

L'absence d'indépendance exprimerait le fait que les émigrants étaient promis, avant même de partir, à une survie supérieure ou inférieure à celle des autochtones, supérieure si la migration porte sur des individus en meilleure santé, inférieure si elle touche des catégories sociales défavorisées, pressées de s'expatrier par des conditions économiques insupportables. Mais de tels mécanismes sont difficiles à illustrer simplement à partir de l'exemple choisi. D'où le recours à un second cas.

Ici, un fichier administratif permet de mesurer à partir de 60 ans la mortalité par âge des ouvriers du bâtiment selon que ceux-ci sont encore actifs ou déjà retraités (tableau 2). Les deux populations exposées au risque évoluent en sens inverse, le groupe des actifs étant progressivement amputé par les départs en retraite, qui viennent gonfler peu à peu l'effectif des retraités (tableau 3, colonnes 1 et 4). Le maximum des départs a lieu à 65 ans, âge légal pour ces générations (1900-1904), mais les départs anticipés ne sont pas négligeables de 60 à 64 ans. Numériquement, il y a trop peu de retraités à 60 ou 61 ans et trop peu d'actifs à partir de 65 ans pour qu'une comparaison de la mortalité des deux groupes puisse porter sur ces âges ; mais c'est possible à 62-64 ans, où le quotient annuel moyen des actifs atteint 21 pour 1 000 contre 49 pour 1 000 pour les retraités (tableau 3, colonnes 3 et 6). Pour ces derniers, le calcul peut se prolonger au-delà, puisque les effectifs ne cessent de croître : à 65-67 ans, le quotient annuel moyen de mortalité n'est plus que de 37 pour 1 000 (tableau 3, colonne 6).

TABLEAU 2. – ÂGE À LA RETRAITE ET ÂGE AU DÉCÈS À PARTIR DE 60 ANS, OUVRIERS DU BÂTIMENT, GÉNÉRATIONS 1900-1904

Age au décès	Décès d'actifs	Décès de retraités suivant l'âge à la retraite									
		60 ans	61	62	63	64	65	66	67	68 +	Total
60 ans	44	6									6
61 ans	49	18	1								19
62 ans	49	31	4	4							39
63 ans	71	21	9	5	6						41
64 ans	77	28	12	18	9	5					72
65 ans	48	27	9	8	9	5	61				119
66 ans	32	26	11	12	11	12	65	1			138
67 ans	27	26	9	11	10	6	77	3	1		143
68 +	5	320	121	177	147	115	2 298	154	81	64	3 477
Total	402	503	176	235	192	143	2 501	158	82	64	4 054

Le rapprochement des deux derniers quotients donne un résultat choquant : la mortalité baisse lorsqu'on avance en âge. Le découpage du groupe des retraités selon l'âge au départ en retraite permet d'en saisir

TABLEAU 3. – CALCUL DES QUOTIENTS DE MORTALITÉ DES ACTIFS ET DES RETRAITÉS

Age	Encore actifs (1)	Décès d'actifs (2)	Quotient de mortalité (3)		Déjà retraités (4)	Décès de retraités (5)	Quotient de mortalité (6)	
60	3 953	44	0,011		503	6	0,012	
61	3 733	49	0,013		673	19	0,028	
62	3 449	49	0,014		889	39	0,044	
63	3 208	71	0,022	0,021	1 042	41	0,039	0,049
64	2 994	77	0,026		1 144	72	0,063	
65	416				3 573	119	0,033	
66	210				3 612	138	0,038	0,037
67	96				3 556	143	0,040	

TABLEAU 4. – CALCUL DES QUOTIENTS DE MORTALITÉ DES OUVRIERS PRENANT LEUR RETRAITE À 60 ANS ET À 65 ANS

Age	Retraite à 60 ans				Retraite à 65 ans			
	Retraités	Décès	Quotient de mortalité		Retraités	Décès	Quotient de mortalité	
60	503	6	0,012					
61	497	18	0,036					
62	479	31	0,065					
63	448	21	0,047	0,060				
64	427	28	0,066					
65	399	27	0,068		2 501	61	0,024	
66	372	26	0,070	0,071	2 440	65	0,027	0,028
67	346	26	0,075		2 375	77	0,032	

l'origine. A âge égal (65-67 ans), ceux qui ont pris leur retraite à 65 ans ont une mortalité beaucoup plus faible que ceux qui ont connu un départ anticipé dès 60 ans (28 pour 1 000 contre 71 pour 1 000) (tableau 4). Pour ces derniers, l'évolution avec l'âge est normale : la mortalité à 65-67 ans est plus forte qu'à 62-64 ans (71 pour 1 000 contre 60 pour 1 000). Lorsqu'on confond dans un calcul unique la mortalité par âge de tous les retraités, on adjoint entre 64 et 65 ans un groupe de retraités partis à l'âge légal à un groupe de « retraités par anticipation ». La composition de la population au risque se trouve ainsi sensiblement modifiée.

L'âge au départ en retraite n'est vraisemblablement pas en cause, mais les départs aux âges les plus précoces sont le fait de populations fragiles, voire invalides, ayant obtenu une retraite anticipée. Ceux qui attendent 65 ans sont en meilleure santé. Imaginons que la législation ait été modifiée au moment où la génération atteignait 60 ans, rendant possible pour tous une cessation d'activité dès cet âge dans de bonnes conditions financières. Les personnes qui avaient initialement fait le projet d'attendre 65 ans partiront alors cinq ans plus tôt, mais comment penser que la mortalité de ces retraités sera, à 62-64 ans, aussi forte que celle des ouvriers qui avaient prévu de partir dès 60 ans, souvent pour des raisons de santé ? La condition d'indépendance n'est donc pas vérifiée ici, puisque les personnes qui prennent leur retraite à 65 ans n'auraient sans doute pas eu, à 62-64 ans, la même mortalité que leurs collègues partis avant l'âge légal, s'ils avaient pu avancer leur propre départ.

L'absence d'indépendance trouve son origine dans *l'hétérogénéité* des conditions de santé au sein des ouvriers du bâtiment et dans l'*association* qui en résulte entre l'âge au départ en retraite et les risques de mortalité à chaque âge. Une simple représentation graphique suggérant que les quotients de mortalité reculent entre 62-64 et 65-67 ans serait fautive, *a fortiori*, une combinaison de ces quotients dans une table synthétisant la mortalité des retraités. Reprenons les ordres de grandeur observés chez les ouvriers du bâtiment, pour illustrer les effets de l'hétérogénéité et de l'association entre âge au départ en retraite et mortalité :

— les ouvriers en bonne santé ont des quotients de mortalité de 21 et 28 pour 1 000 à 62-64 et 65-67 ans ; pour ceux en mauvaise santé, ce sont respectivement 60 et 71 pour 1 000 ;

— 70 % des ouvriers cessant leur activité avant 65 anssont en mauvaise santé et 30 % en bonne santé ; 100 % des ouvriers prenant leur retraite à 65 ans sont en bonne santé ;

— 1 000 ouvriers partent en retraite avant 65 ans et 2 500 à 65 ans.

A 62-64 ans, le quotient de mortalité des retraités serait ainsi calculé sur environ[2] 1 000 personnes dont 700 en mauvaise santé et 300 en bonne santé, soit 0,7 × 60 pour 1 000 + 0,3 × 21 pour 1 000 = 48 pour 1 000.

A 65-67 ans, il serait calculé sur 2 500 personnes supplémentaires, soit 3 500 personnes environ[3] dont 700 (= 20 %) en mauvaise santé et 2 500 + 300 = 2 800 (= 80 %) en bonne santé ; ce quotient serait donc égal à 0,2 × 72 pour 1 000 + 0,8 × 28 pour 1 000 = 37 pour 1 000. On retrouve ainsi les résultats obtenus précédemment, mais on voit plus clairement sur ces exemples chiffrés que l'incomparabilité des quotients à 62-64 et 65-67 ans tient au changement de composition de la population ; à partir de 65 ans on passe d'une minorité à une majorité de bien portants. Mesurer à l'état pur la mortalité des retraités aux différents âges, c'est se demander comment celle-ci évoluerait, en l'absence des modifications introduites par l'entrée progressive en retraite. On le fera ici en supposant que le groupe des retraités est, dès 60 ans, dans la composition qu'il aurait si tous les ouvriers devaient prendre leur retraite à cet âge (20 % en mauvaise santé et 80 % en bonne santé). On aura ainsi à 62-64 ans : 0,2 × 60 pour 1 000 + 0,8 × 21 pour 1 000 = 28 pour 1 000[4], qui est directement comparable avec le quotient de 37 pour 1 000 à 65-67 ans. La combinaison

(2) Environ, car quelques personnes prennent leur retraite à 62, 63 ou 64 ans et s'adjoignent progressivement au groupe des retraités partis avant eux.

(3) Environ, car il faut retirer le nombre de retraités déjà décédés.

(4) Environ, car le poids des retraités en mauvaise santé diminue plus vite que celui des autres retraités.

28-37 pour 1 000 se substitue ainsi à la combinaison 49-37 pour 1 000, faussée par l'absence d'indépendance.

Que faire en l'absence d'indépendance ?

De ces deux illustrations du principe d'indépendance et de ses conséquences pour l'analyse démographique, on peut tirer deux conclusions, dont la première s'adresse plutôt à D. Courgeau et É. Lelièvre et la seconde plutôt à X. Thierry.

En l'absence d'indépendance entre le phénomène observé et le phénomène perturbateur, celui-ci peut modifier la composition de la population exposée au risque dans des proportions telles que le rapprochement des quotients aux âges successifs n'a plus de sens. L'addition et la répartition par âge de ces quotients ne peuvent alors être des mesures satisfaisantes de l'intensité et du calendrier du phénomène. Dans l'exemple ci-dessus, l'abaissement de l'âge légal à la retraite suffirait à modifier ces deux synthèses, sans que la mortalité ait changé.

Nous pouvons exprimer cette même conclusion très différemment, en partant d'abord d'une situation où l'indépendance est vraisemblable, avant de passer aux conséquences de son absence. Dans le premier cas, on peut combiner les quotients successifs par multiplication de leur complément à l'unité pour obtenir la «*fonction de séjour*», puis calculer les événements de la table de séjour en combinant par multiplication la série des quotients et la fonction de séjour. La somme de ces événements est la mesure classique de l'intensité. On peut obtenir une autre synthèse en appliquant la série des quotients à une population de structure cylindrique, c'est-à-dire d'effectifs constants au fil des âges. Mais le choix de cette structure est arbitraire, comme l'est toujours le choix d'une structure-type dans le calcul de taux standardisés. En l'absence d'indépendance, on ne peut certes obtenir de fonction de séjour, mais la synthèse par une structure-type reste soumise à l'arbitraire du choix de celle-ci : pourquoi pas un cône, qui modifierait peut-être le sens des comparaisons intergroupes auxquelles la méthode est vouée ? Mais plus fondamentalement, quelle signification donner à la synthèse de quotients disparates, parce que calculés sur des populations de composition différente ?

Faut-il abandonner pour autant la recherche d'une synthèse des quotients ? L'exemple des retraités suggère que cette recherche doit suivre une autre voie : on a scindé le groupe en deux sous-groupes plus homogènes, sur la base de leur état de santé, facteur commun à la mortalité (phénomène observé) et à l'âge à la cessation d'activité (événement perturbateur). Supposons de même, sur un exemple fictif, que la population atteignant l'âge nubile était composée, avant la diffusion de la cohabitation, de deux sous-populations : l'une, très minoritaire (10 %), avait de faibles chances de se marier, tandis que l'autre pesait si lourd dans le total (90 %) que ses fortes chances s'écartaient peu de la moyenne (tableau 5A). L'état de santé de ces deux groupes pourrait être, là aussi, le facteur discriminant. Supposons

TABLEAU 5. – MESURE DE LA NUPTIALITÉ AVANT ET APRÈS LA DIFFUSION DE LA COHABITATION

A. Avant	Groupe à forte nuptialité		Groupe à faible nuptialité		Ensemble		
Age	Céliba-taire*	Mariages	Céliba-taire*	Ma-riages	Céliba taire*	Mariages	Quotient de nuptialité
17-19	900	225	100	0	1 000	225	0,225
20-21	675	225	100	5	775	230	0,297
22-24	450	225	95	5	545	230	0,422
25-29	225	112,5	90	10	315	122,5	0,389
30-34	112,5	45	80	0	192,5	45	0,234
35+	67,5		80		147,5		

B. Après	Groupe à forte nuptialité**			Groupe à faible nuptialité		Ensemble			
Age	Hors union*	Mariages	Entrées cohabitat.	Hors union	Ma-riages	Hors union*	Ma-riages	Entrées cohabi-tat.	Quotient de nuptialité
17-19	900	196,9	196,9	100	0	1 000	196,9	196,9	0,218
20-21	506,25	140,6	140,6	100	5	606,25	145,6	140,6	0,272
22-24	225	84,4	84,4	95	5	320	89,4	84,4	0,322
25-29	56,25	21,1	21,1	90	10	146,25	31,1	21,1	0,229
30-34	14,06	4,5	4,5	80	0	94,06	4,5	4,5	0,049
35+	5,06			80		85,06			

* Au début de l'intervalle d'âge.
** Quotients de nuptialité égaux à ceux du cadre A et égaux aux quotients d'entrée en cohabitation.

en outre que la diffusion de la cohabitation ne modifie pas ce clivage mais qu'elle donne aux «plus mariables» une probabilité de former une cohabitation égale à celle de se marier et *indépendante d'elle.* Il en résulterait, pour l'ensemble de la population, un recul des quotients de nuptialité (par mariage direct), bien que ceux-ci soient restés inchangés dans chaque sous-population (tableau 5B). Le mécanisme est le suivant : une forte probabilité d'entrée en cohabitation s'ajoutant à la nuptialité dans la sous-cohorte la plus mariable, celle-ci connaît, au fil des âges, une baisse sensiblement accélérée des effectifs encore hors union ; le poids relatif du groupe dans la nuptialité de l'ensemble est donc sensiblement moindre, aux divers âges, qu'avant la diffusion de la cohabitation.

L'absence d'indépendance entre nuptialité et entrée en cohabitation, rendue patente dans les générations françaises par le recul de la première à mesure que la seconde se développe, pourrait ainsi résulter d'une hétérogénéité des cohortes et d'une association entre les intensités des deux phénomènes. On rétablirait donc une mesure synthétique de la nuptialité des générations en pondérant les fonctions de séjour (proportions de célibataires) des deux sous-populations par le poids initial de celles-ci. Dans l'exemple fictif ci-dessus, on obtient ainsi la même mesure de la nuptialité avant et après la diffusion de la cohabitation.

Une dépendance entre phénomènes observés et phénomènes perturbateurs ne doit donc pas faire inévitablement abandonner le projet de combiner les quotients successifs dans une fonction de séjour. Il faut développer une analyse démographique différentielle, comme *préalable* à la détermination de mesures synthétiques. Cette analyse n'est pas nécessairement une étude des facteurs socio-économiques des phénomènes démographiques, mais une recherche des facteurs qui déterminent l'hétérogénéité des populations, à la fois, à l'égard des phénomènes observés et perturbateurs. Cette conclusion, fondée sur quelques observations empiriques ou quelques constructions hypothétiques, se contente de reproduire celle qu'atteignait L. Henry par une analyse systématique des effets de l'hétérogénéité des cohortes et de l'association entre phénomènes démographiques sur la mesure de ceux-ci[5].

Patrick FESTY

[5] Louis Henry, « D'un problème fondamental de l'analyse démographique », *Population*, 1, 1959.