

# JOURNAL DE LA SOCIÉTÉ STATISTIQUE DE PARIS

ROLAND PRESSAT

## **Le caractère unitaire des méthodes de la démographie**

*Journal de la société statistique de Paris*, tome 125, n° 3 (1984), p. 151-157

[http://www.numdam.org/item?id=JSFS\\_1984\\_\\_125\\_3\\_151\\_0](http://www.numdam.org/item?id=JSFS_1984__125_3_151_0)

© Société de statistique de Paris, 1984, tous droits réservés.

L'accès aux archives de la revue « Journal de la société statistique de Paris » (<http://publications-sfds.math.cnrs.fr/index.php/J-SFdS>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques  
<http://www.numdam.org/>

## LE CARACTÈRE UNITAIRE DES MÉTHODES DE LA DÉMOGRAPHIE (\*)

Roland PRESSAT

*Chef du département de conjoncture,  
Institut national d'études démographiques*

*L'analyse des statistiques démographiques gagne à être conduite par référence au modèle de base que constitue la table démographique. Ainsi se dégagent les notions d'intensité et de calendrier. Il convient encore de préciser les liens qui existent entre les nombreux taux démographiques et les éléments des tables correspondantes à la fois pour dégager la signification exacte des premiers et mieux comprendre la signification des indices transversaux.*

*Population statistics would benefit by the use of the demographic table which constitutes a basic model. The notions of intensity and timing can be derived from it. We must also thoroughly examine the link between the various demographic rates and the elements from which they are derived. We thus obtain a better understanding of the real significance of the rates as well as that of transversal indices.*

L'analyse des statistiques démographiques a été pendant longtemps et est encore souvent teintée du plus large empirisme, les indices couramment utilisés formant un ensemble disparate présentant tout au plus quelque attache avec la notion de probabilité.

Pendant longtemps, la référence aux méthodes utilisées pour étudier la mortalité s'est imposée : l'âge apparaissait alors comme une variable essentielle pour décrire les phénomènes démographiques et l'on accordait aux divers taux (par âge) une signification analogue à celle des taux de mortalité.

Un effort dans le sens d'une unification des méthodes fondée sur une analyse plus pertinente de la signification des diverses fréquences que calcule le démographe, a été accompli au cours des quarante dernières années. Sans doute trouve-t-on encore des approches empiriques, spécialement dans la littérature anglo-saxonne. Et c'est peut-être l'opposition entre cette volonté unificatrice qui s'est fait jour dans la démographie française, et la persistance d'une présentation pragmatique présente dans les travaux et plus encore dans les manuels de langue anglaise qui fait parfois parler d'une école française de démographie qui se distinguerait de l'école anglo-saxonne. Ajoutons que l'extension considérable des travaux d'inspiration américaine concernant le Tiers Monde où fleurissent des méthodes dites parfois « non conventionnelles », n'a pu que creuser le fossé entre une démographie à prétention unitaire quant à ses méthodes et ce qui est avant tout un arsenal de recettes.

Cette distinction que nous opérons n'entend pas pour autant discréditer telle ou telle recherche démographique, dès lors que celle-ci est menée avec la rigueur qui est de mise. Elle vise plutôt à montrer le bénéfice que l'on peut retirer, au plan didactique notamment, d'une vision plus solidement construite de ce que sont les méthodes de l'analyse démographique.

---

(\*) Communication faite le 21 mars 1984 devant les Sociétés de statistique de Paris et de France.

## LE MODÈLE DE BASE : LA TABLE DÉMOGRAPHIQUE

Le démographe est confronté à des données brutes fournies par les relevés d'état civil, les recensements, les enquêtes et des relevés spécifiques. Il a pour tâche d'en tirer des produits élaborés, significatifs des comportements humains et se prêtant mieux que la matière première dont ils sont issus, à l'interprétation et à l'analyse causale.

Graunt a donné l'exemple le premier, il y a plus de trois siècles, en osant proposer une analyse statistique de la mortalité qui l'amenait à conclure à une certaine permanence de ses manifestations.

La table de mortalité est de fait l'ancêtre des produits finis que le démographe moderne est appelé à proposer pour conduire à son terme l'analyse des phénomènes de population. Mais c'est, disions-nous, un asservissement aveugle à ce modèle descriptif qui est répréhensible; et cette critique nous amène à présenter d'entrée de jeu la généralisation correcte que l'on peut faire de la notion de table démographique et à montrer la base solide que constitue cette notion en tant qu'outil d'analyse.

Rappelons ici quelques points d'histoire et rendons hommage à l'ancien Secrétaire général de la Société de statistique de Paris, Pierre Depoid qui, comme nous aurons l'occasion de le voir, a été dans plusieurs domaines de la démographie quantitative un novateur. C'est véritablement sous sa plume que pour la première fois on rencontre, en France du moins, le premier élargissement correct de la notion de table de mortalité avec la présentation sous sa forme moderne, de la table de nuptialité : dans le Bulletin de la Statistique générale de la France de janvier-mars 1938, Depoid présentait des « Tables nouvelles relatives à la population française », au nombre desquelles figurait la table de nuptialité des célibataires pour la période 1930-1932. Sans doute, sous ce nom, une table avait été publiée par Michel Huber, un ancien Président de la Société de statistique et Directeur de la Statistique générale. L'apport de Pierre Depoid consistait à définir une table *en l'absence de mortalité*, donc une table qui entendait donner une mesure de la nuptialité à l'état pur. On peut être surpris que cette notion ait mis aussi longtemps à voir le jour alors que le taux brut de reproduction, qui mesure la fécondité en l'absence de mortalité, était un concept en vigueur depuis le XIX<sup>e</sup> siècle... Il appartenait enfin à Paul Vincent qui, lui aussi, s'est montré très actif au sein de la Société de statistique de Paris, d'apporter une correction mineure aux calculs de Depoid pour que la table de nuptialité voit le jour, dans sa forme à peu près actuelle. On trouvera ce dernier travail en annexe à l'article de l'auteur intitulé « La famille normale » paru dans le numéro d'avril-juin 1950 de *Population*. On peut seulement noter que le contenu précis d'une table n'est pas encore codifié : figurent dans la table de Depoid, revue par Vincent, la série des quotients et celle des célibataires mais non celle des mariages.

On jugera de la lenteur avec laquelle cette notion de table démographique se dégage des balbutiements empiriques en parcourant la littérature démographique contemporaine. C'est ainsi que dans leur ouvrage « *The Methods and Materials of Demography* », ouvrage de base en langue anglaise, Shryock et Siegel présentent sous l'intitulé « Table de nuptialité pour les femmes blanches aux États-Unis : 1958 à 1960 », un ensemble complexe de colonnes aboutissant à la présentation des effets conjoints de la mortalité et de la nuptialité sur l'apparition des nouveaux mariages. On débouche ainsi sur ce que l'on peut appeler une table de nuptialité *nette* laquelle revient à reconstituer le réel, alors qu'une authentique démarche analytique consiste, au contraire, à partir de la réalité observée pour essayer d'en séparer les constituants, soit d'une part, les caractéristiques *pures* de la nuptialité des célibataires, c'est-à-dire sans que s'interpose la mortalité, et d'autre part, les caractéristiques de la mortalité des célibataires. On trouve encore dans la littérature démographique d'autres exemples de construction de portée limitée tel ce monstre hybride qu'est la *table de vie active*.

Mais l'ambiguïté qui subsiste quant à ce que doit être une authentique table démographique ne concerne pas que la substance de celle-ci, mais encore l'échelle temporelle sur laquelle doit s'appuyer la description.

La référence implicite à la mortalité a fait jouer pendant longtemps un rôle exclusif à l'âge en tant que support descriptif. Le divorce offre un exemple évident où l'âge n'intervient, en quelque sorte, qu'au deuxième degré, la durée de l'union rompue étant le facteur temporel primordial. Dès 1952, Louis Henry insistait dans son étude intitulée « Mesure de la fréquence des divorces » sur le bien-fondé de ce choix, sans écarter naturellement l'intérêt que peut présenter une prise en compte simultanée de la durée du mariage et de l'âge des personnes concernées, ce qui revient à faire intervenir l'âge des personnes au moment de la conclusion de l'union en cause.

Plus systématiquement, Henry posait comme principe d'étudier la venue des événements relatifs à un phénomène donné comme consécutive à la venue d'un événement antérieur qui permet la surveillance des événements caractéristiques du phénomène étudié. C'est ainsi que le divorce sera étudié comme consécutif au mariage, le remariage comme consécutif au divorce ou au veuvage, la venue d'une naissance d'un rang donné comme consécutive à la venue de la naissance de rang précédent, etc. On peut s'étonner que des principes aussi évidents n'aient pas été appliqués depuis longtemps et que, maintenant encore, ils ne débouchent pas sur une mise en œuvre systématique. La cause nous paraît être double : l'absence de statistiques adéquates et une certaine paresse mentale qui occulte toute pensée critique face à certaines démarches formelles.

#### TABLE LONGITUDINALE, TABLE TRANSVERSALE

Les tables démographiques, plus ou moins correctement établies, ont vu le jour dans une optique transversale. Et ce n'est vraiment qu'au cours des quarante dernières années et après certains travaux pionniers dont ceux de Delaporte sur la mortalité et de Depoid encore sur la fécondité, que l'on a reconnu, non seulement la prééminence des analyses longitudinales, mais encore l'ambiguïté des synthèses transversales. Il n'y a guère que la mortalité qui échappe à cette critique des tables transversales dans la mesure où la situation d'une année semble peu redevable aux conditions des années précédentes. Au contraire, les solidarités qui se manifestent au long de la vie des individus quant aux modalités selon lesquelles ils se marient et procréent font que le report, avec comme support une génération ou une cohorte, d'une suite d'histoires partielles empruntées à des cohortes différentes, s'apparente à une fiction.

#### INTENSITÉ ET CALENDRIER

Dans leur forme, les tables démographiques correctement établies se différencient selon qu'elles entendent décrire des phénomènes à événements non renouvelables (tables de mortalité, de nuptialité des célibataires, de divortialité, de fécondité ou de migration de rang) ou à événements renouvelables (tables de fécondité sans distinction de rang). Les premières appelées encore tables à extinction (exactement des tables à *simple* extinction) rassemblent, pour une suite correctement choisie de durées (âges, durées de mariage, ...), une colonne de survivants, une colonne d'événements et une colonne de quotients; les secondes ne réunissent qu'une colonne d'événements et une colonne d'événements cumulés.

Ce qui fait l'unité de ce mode de description, c'est que, dans un cas comme dans l'autre, il fait apparaître ces deux grandeurs fondamentales de l'analyse démographique que sont le *calendrier* des phénomènes et leur *intensité*, c'est-à-dire la façon dont se répartissent les événements au long de la vie des cohortes envisagées, et la fréquence des événements par tête dans ces cohortes. Cela est possible dans les deux types de table car, dans un cas comme dans l'autre, on s'attache précisément à faire apparaître la séquence des événements telle qu'elle s'établit au long de la vie de la cohorte.

On a pu faire la critique de ces deux notions d'intensité et de calendrier en tant qu'outil d'analyse, dans la mesure où elles ne sont pas indépendantes. C'est sans doute une marque de leur insuffisance, mais cela ne les condamne pas pour autant; c'est seulement une invite à être nuancé et prudent dans les interprétations s'appuyant sur elles.

#### TAUX DÉMOGRAPHIQUES ET TABLES DÉMOGRAPHIQUES

Le cadre conceptuel des tables démographiques longitudinales (les seules ayant un sens non équivoque) constitue une référence permanente plus ou moins explicite dans toutes les analyses démographiques. Il reste alors à voir comment, à partir des données d'observation, on peut élaborer des histoires statistiques plus ou moins étendues dans les formes prévues par les tables.

Les indices fondamentaux de l'analyse démographique sont les taux. Ces indices se sont imposés de façon empirique : en éliminant ce que la fréquence des événements doit à l'importance des populations, les taux constituent une première base de comparaison et par conséquent d'analyse. Toutefois, la confusion est grande quant à la façon de les calculer et à la signification à leur accorder, le même mot recouvrant des entités de sens très variés. C'est cet empirisme qui a présidé à leur construction qui les a longtemps fait considérer comme étant des fins en eux-mêmes sans que leur signification profonde ait été dégagée.

Il est évidemment regrettable que le mot taux désigne, selon les cas, des indices de signification très différente. Nous sommes ici les héritiers du long passé en friches dans lequel a vécu la démographie comme science quantitative. On s'est parfois élevé contre cet emploi abusif d'un mot appliqué à des indices de nature différente. C'est ainsi que L. Henry, se référant à la signification revêtue par une certaine classe de taux, a forgé l'expression « événements réduits ». A suivre strictement ce nouveau mode d'appellation, on devrait bannir du vocabulaire démographique l'expression « taux de fécondité ». Pareillement, et en raison d'un autre type de distinction, taux de scolarité et taux d'activité devraient céder la place à proportions de scolarisées, proportions d'actifs. Nous voyons mal comment cette petite révolution du vocabulaire pourrait s'accomplir et c'est pourquoi nous restons attaché à l'appellation taux, assortie de qualificatifs qui marquent les différences. Nous avons ainsi parlé de taux de première catégorie et taux de deuxième catégorie. Les premiers, dont le type est le taux de mortalité, ont une parenté étroite avec les quotients de la table attachée au phénomène considéré; les seconds, dont le type est le taux de fécondité (naissances de tous rangs réunis), s'identifie aux événements de la table : ce sont précisément ceux que L. Henry dénomme événements réduits.

Dans l'un et l'autre cas, il s'agit de fréquences d'événements dans des populations. Et c'est le choix des populations devant figurer au dénominateur qui fait la différence. D'une composition toute différente sont les taux du troisième type que, dans notre timidité, nous n'avons pas osé dénommer taux de troisième catégorie : il s'agit cette fois de rapports de sous-populations à populations. Le type en est le taux de célibat, rapport dans une population, à une date donnée et à un âge donné, des célibataires à l'ensemble de la population du même âge. Ce taux s'identifie sensiblement à la série des célibataires de la table de nuptialité correspondante, c'est-à-dire, en toute généralité, à la série des survivants. Rappelons à cet égard que c'est le statisticien anglais Hajnal qui a proposé le premier vers la fin des années quarante (après toutefois des tentatives dans ce sens faites par Mortara au Brésil) en la justifiant, l'assimilation des taux de célibat aux célibataires de la table. Il est significatif que personne n'ait vu, pendant longtemps, que ce que l'on appelle encore la méthode de Hajnal n'est qu'un cas particulier de l'utilisation que l'on peut faire de ces taux relevant de la troisième catégorie.

Ainsi donc un pont est jeté entre taux et éléments des tables démographiques. Et l'interprétation à donner aux premiers doit toujours se faire par référence à leurs correspondants dans ces tables. Mieux

encore, ces taux sont, le plus souvent, l'intermédiaire commode qui permet de déterminer l'ensemble des éléments constitutifs d'une table.

#### DIFFICULTÉS DES MESURES TRANSVERSALES

Le moment est maintenant venu d'étudier comment la parenté que nous reconnaissons entre les diverses classes de taux et les classes correspondantes d'éléments des tables trouvera son utilité en analyse transversale.

Les tables démographiques sont d'abord apparues dans une optique transversale. Comme nous l'avons rappelé, c'est à Delaporte et Depoid que nous devons les premiers essais de reconstruction longitudinale à partir d'une suite d'observations par période. En transversal, ce sont les quotients qui introduisent à la table, du moins pour ce qui est des tables à extinction, celles pour lesquelles la notion de quotients existe; s'agissant des tables à événements renouvelables, c'est le recours aux taux de deuxième catégorie qui constitue la voie d'accès. C'est en travaillant à partir de ces taux, mais réduits aux seules naissances de premier rang, qu'au cours des années quarante, le démographe américain Whelpton rencontra un paradoxe : la somme des taux relatifs à une année faisait apparaître plus d'une naissance de premier rang par femme. Ce paradoxe, comme souvent dans l'histoire des sciences, fut un révélateur. En cherchant à le résoudre, Whelpton mit le doigt sur l'importance que revêtent les fantaisies des calendriers des diverses cohortes dans les manifestations d'un phénomène une année donnée. Il y a fort à parier que cette inconsistance formelle incita l'auteur à porter ses efforts sur les reconstitutions longitudinales par lesquelles il s'est rendu célèbre, et avec lesquelles il ne risquait pas de rencontrer ce qui a toutes les apparences d'une incongruité.

Whelpton avait ainsi calculé, avant la lettre, ce que L. Henry dénomme la somme des naissances réduites de premier rang; nous préférons parler d'indice synthétique de fécondité de premier rang.

Tous les phénomènes sont susceptibles d'être synthétisés en transversal par un « indice synthétique » (toutefois en matière de mortalité, l'occasion est plus exceptionnelle et l'intérêt moindre). Et, bien loin de nous inciter à renoncer à son calcul, le fait que les résultats auxquels il conduit ont toutes les apparences d'une extravagance, doit être soigneusement analysé et mis à profit au plan de l'interprétation. En fait, les indices synthétiques construits comme somme des taux de deuxième catégorie (une somme des événements réduits, dit L. Henry) est l'indice transversal le plus simple d'interprétation parce que c'est celui qui se prête le plus commodément à l'analyse de sa composition. Un indice synthétique est une fonction d'éléments de calendriers et des intensités des cohortes qui participent à sa détermination, ce que ne doit pas faire oublier le fait que ledit indice s'exprime en nombre moyen d'événements par tête (ainsi un nombre moyen de naissances vivantes par femme).

On jugera de la confusion méthodologique qui a longtemps prévalu dans le fait qu'on a bâti (et que l'on bâtit encore parfois) des tables démographiques du moment selon, nécessairement, deux principes différents; ainsi :

- une table de mortalité du moment repose sur la mise en jeu d'une série de quotients;
- une table de fécondité du moment (tous rangs de naissance confondus) ne saurait reposer que sur l'addition de la suite des événements (les taux de fécondité).

On mesure bien la différence qui peut exister entre ces deux modes de confection de l'indice du moment sur l'exemple de la nuptialité des célibataires où les deux types de calcul sont possibles. Mais ce qui est décisif dans le choix à opérer, quand un tel choix est possible, c'est la relative facilité d'analyse de la composition de l'indice synthétique face à cette difficulté d'interprétation des données d'une table du moment, difficulté due à la complexité du mode d'enchaînement des caractéristiques élémentaires qui entrent en jeu.

L'analyse transversale s'efforce de saisir une situation qui est le plus souvent en perpétuel mouvement. Et c'est évidemment une imprudence que d'interpréter en termes d'histoires vraies ce qui n'est qu'une mise, bout à bout, d'une suite de tronçons d'histoire appartenant à des cohortes différentes. La démographie a longtemps vécu sur cette fiction et il a fallu les bouleversements causés par la seconde guerre mondiale, la crise économique qui l'a précédée, et un bouleversement de certains comportements (en matière de fécondité essentiellement) pour que l'on prenne conscience des limites des synthèses transversales. Ces dernières, quand elles sont bien choisies, sont au plus des indicateurs (on a d'ailleurs proposé d'appeler indicateur conjoncturel l'indice synthétique de fécondité), des sortes de baromètres. Mais les constituants de ces indicateurs (ainsi les taux de fécondité générale par âge) sont à replacer dans l'histoire des cohortes en cause et, finalement, l'analyse transversale proprement dite repose essentiellement sur le commentaire que permet cette adjonction des données de l'année ou de la période analysée, à l'ensemble des positions acquises par les diverses cohortes.

Il n'est pas jusqu'à l'analyse de la mortalité qui ne se porte mieux d'une vision par cohorte. C'est vrai particulièrement lorsque l'analyse porte sur une cause de décès en évolution rapide. Deux cas classiques se trouvent dans la littérature épidémiologique :

- le cancer broncho-pulmonaire dont on voit la fréquence augmenter de génération en génération et, au sein d'une même génération, avec l'âge de cette dernière; une vision transversale, si elle suggère bien une augmentation dans le temps à tous les âges, suggère une diminution du risque avec l'âge au-delà de 70 ans : pure illusion, cette diminution tient au fait qu'on l'observe précisément dans des générations différentes, dont les histoires en matière de comportement (nous pensons ici essentiellement à la consommation de cigarettes), ont sensiblement varié dans le temps. Ainsi donc, on ne doit pas conclure à une diminution du risque avec l'âge, mais bien à son aggravation constante;
- la mortalité par tuberculose pulmonaire qui, à l'inverse, a décréu de génération en génération, tout en restant une cause de décès qui frappe préférentiellement les personnes en début d'âge adulte. Vue année par année, l'évolution observée suggère que l'âge au risque maximal se déplace progressivement vers des âges de plus en plus élevés alors qu'ici encore l'explication réside dans l'histoire passée des individus (moindres soins efficaces dans le passé, séquelles d'affections anciennes).

A la limite, l'analyse transversale peut apparaître comme une tâche impossible sinon rétrospectivement une fois connue l'histoire complète des cohortes impliquées durant les années considérées. Mais alors l'analyse ne remplit plus son rôle qui est de suivre au mieux l'actualité. Cela ne doit pas dissuader d'entreprendre une telle tâche au fil des années, mais retenons la prudence avec laquelle il convient de conclure et la nécessité de s'en remettre non pas seulement à quelques mesures synthétiques mais encore à des rapprochements variés au sein des cohortes en cause.

#### DIGRESSION SUR LES MESURES DE REPRODUCTION

L'analyse transversale, ou du moins l'usage que l'on a fait de certains indices transversaux, est à l'origine de mécomptes particulièrement dans l'étude de la reproduction de la population. Mais l'analyse longitudinale elle-même n'est pas sans équivoque. C'est ici qu'apparaît encore le rôle de précurseur tenu par P. Depoid. Depoid constata, au cours de ses travaux sur la reproduction en Europe, que les taux de reproduction nette à la naissance et à 15 ans ne donnaient pas les mêmes résultats, cela en raison de la baisse de la mortalité. Dans la seconde mesure à la mortalité de 0 à 15 ans des mères, on substituait dans le calcul celle de leurs filles, ce qui amenait à prendre en compte le recul de la mortalité dans cette période de la vie accomplie au cours d'une période de 25 à 30 ans. D'autres esprits perspicaces ont

