JOURNAL DE LA SOCIÉTÉ STATISTIQUE DE PARIS

PAUL DAMIANI

Étude des distributions départementales de la fécondité

Journal de la société statistique de Paris, tome 131, n° 2 (1990), p. 68-78

<http://www.numdam.org/item?id=JSFS_1990__131_2_68_0>

© Société de statistique de Paris, 1990, 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/

ÉTUDE DES DISTRIBUTIONS DÉPARTEMENTALES DE LA FÉCONDITÉ

Paul DAMIANI
INSEE, ancien secrétaire général des Sociétés de Statistique
Hélène MASSÉ
INSERM

Dans cet article, on a analysé les distributions départementales d'un indice de fécondité à différents recensements, entre 1861 et 1982. On a constaté que ces distributions étaient la somme de deux distributions normales. On en a déduit l'existence de deux populations caractérisées l'une par une fécondité de jeunesse, l'autre par une fécondité d'âge mûr. A partir de ces résultats, on a défini un indicateur de tendance de la fécondité lié à la probabilité de coïncidence des structures expérimentale et théorique, de la population.

In this paper, we analyse the distributions by French "departement" of an index of fertility, at different census from 1861 to 1982. We find that they are sum of two normal distributions. We infer that there are two populations with a fertility of youth for the one and a fertility of older for the other. From these results, we define an index of evolution of fertility linked with the probability of coïncidence between experimental and theoretical structures of population.

INTRODUCTION

Dans cette étude, nous avons essayé d'ajuster des lois de probabilité connues sur les distributions départementales d'un indicateur de la fécondité.

Dans un travail précédent [1], nous avions réalisé un tel ajustement sur les taux de mortalité. Nous avions constaté que la distribution des valeurs départementales de ces taux était la somme de deux lois log-normales. Nous montrerons, dans cet article, que la distribution de l'indicateur de fécondité étudié peut être considéré comme la somme de deux lois normales, faisant ainsi apparaître l'existence de deux populations de fécondité différente. Nous essaierons d'expliquer, ensuite, ce résultat. Nous en déduirons, enfin, un indicateur d'évolution de la fécondité.

Définitions

Les indicateurs de fécondité sont calculés par l'INSEE (Institut national de la statistique et des études économiques) à partir des effectifs par âge des femmes, obtenus à chaque recensement, et du nombre de naissances vivantes au cours d'une période, de 3 ans en général, entourant ce recensement. Il s'agit des naissances vivantes comprenant les enfants déclarés vivants et les enfants nés vivants mais morts avant la déclaration à l'état-civil; ces naissances sont classées suivant le lieu de domicile de la mère. Des valeurs de ces indicateurs sont estimées annuellement, entre chaque recensement.

Les principaux indicateurs calculés sont les suivantes :

- *taux de fécondité par âge* : rapport du nombre de naissances vivantes issues des femmes d'un groupe d'âge donné à l'effectif de ces femmes,
- *somme des naissances réduites* : c'est le nombre d'enfants que mettraient au monde, au cours de leur vie, une génération de femmes, non soumises à l'effet de la mortalité, qui aurait à chaque âge la fécondité observée pendant la période considérée. On obtient cet indicateur en faisant la somme des taux de fécondité par âge,
- *taux brut de reproduction* : c'est le nombre moyen de filles par femme. Cet indicateur s'obtient en multipliant le précédent par le taux de féminité des naissances, c'est-à-dire la proportion des naissances du sexe féminin par rapport à l'ensemble des naissances. Ce taux reste pratiquement constant dans le temps [2]; il est pris égal à 0,488,
- *taux net de reproduction* : c'est le nombre de filles que mettraient au monde, au cours de leur vie, une génération de femmes ayant à chaque âge la mortalité et la fécondité observées pendant la période considérée,
- *nombre moyen d'enfants par femme* : s'obtient en appliquant, à une génération fictive de femmes, les taux de fécondité d'une période considérée. Deux calculs sont possibles suivant qu'on tient compte ou non de la mortalité. Dans le cas où la génération de femmes n'est pas supposée soumise à la mortalité, on retrouve la « somme des naissances réduites » définie précédemment.

Données de base

Nous avons analysé, dans cet article, les distributions départementales de la somme des naissances réduites. Les valeurs de cet indicateur sont publiées par l'INSEE à l'occasion de chaque recensement de 1946 à 1982 [3]. Pour la période antérieure, nous avons utilisé les évaluations de P. Depoid pour les années 1861, 1891, 1911, 1931 [4].

Analyse des distributions départementales

Nous disposons, pour une date donnée, de la distribution D des valeurs départementales de la somme des naissances réduites, notée y.

Nous constatons que l'ajustement d'une loi simple de probabilité sur la distribution D n'est pas satisfaisant. Nous obtenons un meilleur résultat en considérant que cette distribution est la somme de deux distributions suivant chacune une loi normale, D_1 et D_2. Nous appellerons D_1 la distribution de moyenne m_1 la plus faible et D_2 celle de moyenne m_2 la plus forte.

L'ajustement se fait par une méthode d'approximations successives (voir annexe 1). Nous déterminons, ainsi, l'importance relative et les valeurs caractéristiques des distributions composantes D_1 et D_2.

Nous en déduisons que la population générale P se décompose en deux populations P_1 et P_2, pour chacune desquelles la distribution de la somme des naissances réduites suit une loi normale. La répartition des effectifs des populations P_1 et P_2 par rapport à la population générale ainsi que les valeurs moyennes de la fécondité de ces populations sont celles trouvées pour les distributions D_1 et D_2.

Résultats. Il convient de signaler que les résultats trouvés doivent être considérés comme des approximations, par suite des erreurs d'observation et de la méthode employée.

1. Le tableau I fournit l'évolution de 1861 à 1982 de la proportion θ_1 des effectifs de D_1 par rapport à D. Il indique également les valeurs des moyennes m_1, m_2 et des écart-types s_1, s_2 des distributions D_1 et D_2, ainsi que la moyenne y de D.

   La proportion θ_1 augmente de 1861 à 1931 et passe de 0,55 à 0,65. Après la 2e guerre mondiale, elle part d'une valeur plus faible égale à 0,49, en 1946, pour atteindre 0,55 en 1982; mais cette croissance n'est pas régulière, on observe, en effet, deux maximums en 1954 ($\theta_1 = 0,56$) et en 1968 ($\theta_1 = 0,55$).

   Les variations dans le temps des moyennes de D_1 et D_2 suivent celles de la moyenne de D. Les valeurs de m_1 et m_2 représentent respectivement, en moyenne, 94% et 111% de la valeur de y.

   Les graphiques 1 et 2 illustrent l'évolution dans le temps, respectivement, de la moyenne de la distribution D et de l'importance relative θ_1 de la distribution D_1.

2. Le tableau II est relatif au recensement de 1982. Il donne l'importance relative des effectifs estimés, par rapport à l'effectif total, de chaque classe de valeur de y, pour les distributions D_1, D_2 et D. Il fournit également, pour la classe j, l'importance relative θ_{1j} de l'effectif de D_1 par rapport à l'effectif total de la classe.

TABLEAU I

Evolution dans le temps des caractéristiques des distributions D_1, D_2 et D.

Date	Distribution D_1		Distribution D_2		Distribution D	Importance relative de D_1
	Moyenne m_1	Ecart-type s_1	Moyenne m_2	Ecart-type s_2	Moyenne y	θ_1
1861	3,09	0,33	3,79	0,31	3,46	0,55
1891	2,73	0,39	3,34	0,52	3,03	0,59
1911	2,43	0,36	2,84	0,35	2,54	0,60
1931	2,24	0,29	2,72	0,20	2,42	0,65
1946	2,78	0,24	3,34	0,28	2,97	0,49
1954	2,54	0,22	3,12	0,17	2,69	0,56
1962	2,72	0,19	3,16	0,22	2,83	0,53
1968	2,48	0,22	2,89	0,18	2,59	0,55
1975	1,90	0,17	2,04	0,19	1,95	0,54
1982	1,75	0,11	1,99	0,09	1,88	0,55

TABLEAU II

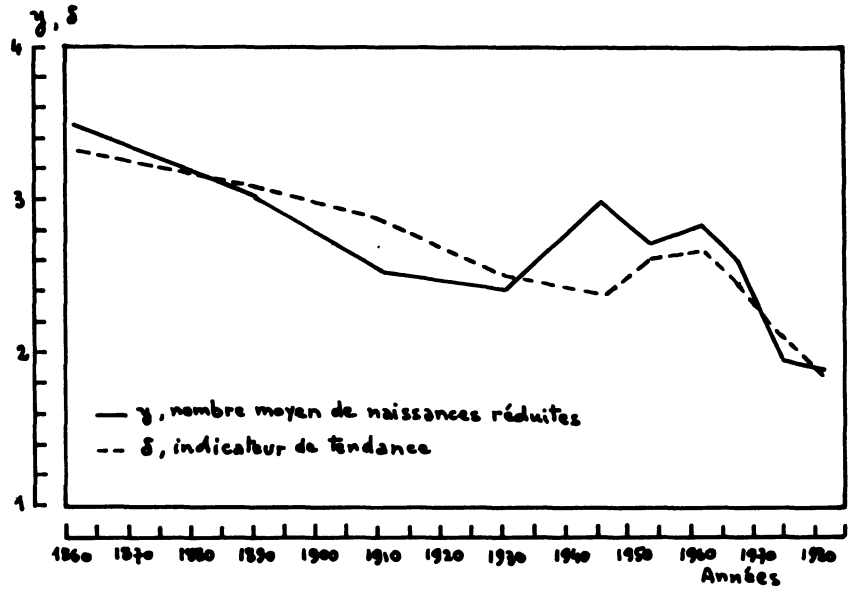
Fréquences estimées des distributions D, D_1, D_2 et importance relative de D_1, suivant la fécondité, en 1982.

Somme des naissances réduites	Fréquences estimées			Importance relative de D_1
	D_1	D_2	D	
y_j	f_{1j}	f_{2j}	f_j	θ_{1j}
Moins de 1,60	0,0495	ε	0,0495	1
1,60 – 1,69	0,1302	0,0002	0,1304	0,998
1,70 – 1,79	0,1896	0,0067	0,1963	0,966
1,80 – 1,89	0,1307	0,0603	0,1910	0,684
1,90 – 1,99	0,0429	0,1726	0,2155	0,199
2,00 – 2,09	0,0066	0,1589	0,1655	0,040
2,10 – 2,19	0,0005	0,0468	0,0473	0,011
2,20 – 2,29	ε	0,0044	0,0044	ε
2,30 – 2,39	ε	0,0001	0,0001	ε
Ensemble	0,5500	0,4500	1	0,550

ε : valeur inférieure à 5.10^{-5}.

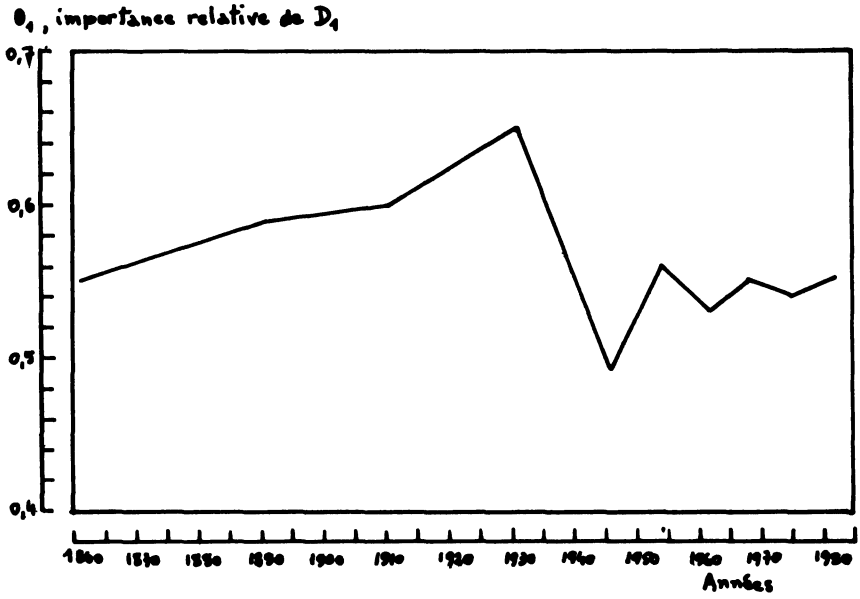
GRAPHIQUE 1

Evolution du nombre moyen de naissances réduites et de l'indicateur de tendance



GRAPHIQUE 2

Evolution de l'importance relative de la distribution D_1



Interprétation des résultats

Pour essayer d'expliquer la signification des deux distributions trouvées, nous avons cherché les facteurs présentant une forte liaison statistique avec la fécondité.

De nombreuses études ont mesuré l'action de différents indices tels que : urbanisation, milieu social, mesures biométriques, etc [5]. Parmi ces indices, nous avons retenu des indicateurs de structure de population, dont l'action paraît la plus prononcée.

Indices de Easterlin et Leridon

H. Leridon a signalé [6] les travaux d'un économiste américain R. Easterlin qui a proposé l'indice suivant : rapport de l'effectif du groupe d'âge 35-64 ans à celui du groupe d'âge 15-34 ans de la population masculine. Ces groupes d'âge caractérisent respectivement le flux de sortie et le flux d'entrée sur le marché du travail. L'évolution de cet indice, pour la période 1940-1976, est comparable à celle du nombre de naissances réduites pour les Etats-Unis et l'Australie. Les résultats sont moins bons pour le Canada et l'Angleterre.

H. Leridon a vérifié cette concordance pour la France, pour la période 1946-1976 (coefficient de corrélation égal à 0,92). Il a proposé un autre indice : rapport de l'effectif du groupe d'âge 60-64 ans à la moitié de l'effectif du groupe d'âge 15-24 ans de la population masculine. Il a trouvé que les variations de cet indice coïncident avec celles du nombre de naissances réduites mais seulement à partir de 1958.

De notre côté, nous avons constaté que les corrélations entre ces deux indices et le nombre de naissances réduites, calculées sur les données régionales de 1982, étaient fortes et négatives dans les deux cas ($r = -0,81$ pour l'indice Easterlin, $r = -0,70$ sur l'indice Leridon).

Autres indices de structure proposés

Nous avons essayé, plus simplement, comme indices de structure, les proportions de population des différents groupes d'âge, pour chaque sexe et pour les deux sexes réunis. La liaison statistique a été mesurée sur les données régionales de 1982.

Les corrélations les plus fortes avec y ont été trouvées pour les deux indices suivants : p_{01} et p_{02}, proportions observées des groupes d'âge 45-64 ans et 0-19 ans pour la population totale.

Le coefficient de corrélation est négatif avec $p_{01}(r = -0,82)$ et positif avec $p_{02}(r = 0,89)$.

Signification des fécondités des populations P_1 et P_2

Nous trouvons également des corrélations fortes entre θ_1 et chacun des indices p_{01} et p_{02}. Le coefficient de corrélation est positif avec $p_{01}(r = 0,78)$ et négatif avec $p_{02}(r = -0,87)$. Autrement dit, la proportion θ_1 de la population P_1 est, respectivement, d'autant plus élevée ou d'autant plus faible que la proportion du groupe d'âge « 45-64 ans » ou « 0-19 ans » est plus forte.

Ces résultats permettent de proposer la signification suivante des fécondités des populations P_1 et P_2. La fécondité de P_2 représenterait surtout celle de la jeunesse, la fécondité de P_1 serait principalement celle de l'âge mûr.

Il convient de remarquer que les groupes d'âge indiqués doivent être considérés comme des valeurs approximatives des groupes d'âge réels correspondant à ces fécondités.

Application : recherche d'un indicateur de tendance

Compte tenu des résultats trouvés, les proportions observées p_{01} et p_{02} des groupes d'âge 45-64 ans et 0-19 ans de la population totale doivent entrer dans la composition d'un indicateur de l'évolution dans le temps de la fécondité.

Nous avons pensé que devaient intervenir également les proportions théoriques p_1 et p_2 de ces mêmes groupes d'âge, calculées à partir des tables de mortalité aux dates considérées (voir annexe 2).

Nous avons ajusté, sur les données de 1861 à 1982, un modèle de régression linéaire entre y et des combinaisons de ces variables. Les conclusions obtenues sont les suivantes :

- il faut éliminer les valeurs correspondant à l'année 1946, qui semblent aberrantes,
- on peut remplacer, sans diminution appréciable de la qualité de l'ajustement, les proportions observées et théoriques par leurs sommes : $p_0 = p_{01} + p_{02}$ et $p = p_1 + p_2$,
- la combinaison de variables la plus simple, qui donne une des meilleures corrélations, est le produit : $u = pp_0$.

  Le modèle de régression obtenu est :

$$y = a + bu + \varepsilon$$

avec :

$$a = -5,1091$$
$$b = 26,9010$$

Le coefficient de corrélation est égal à 0,94.

C'est l'indicateur δ calculé à partir de ce modèle : $\delta = a + bu$, que nous proposons comme *indicateur de tendance de la fécondité.*

Le tableau III donne les variations dans le temps des proportions observées et théoriques, p_0 et p, ainsi que de u et δ. Il s'agit de valeurs arrondies

TABLEAU III

Evolution dans le temps des proportions observées et théoriques des groupes d'âge retenus et de l'indicateur de tendance

Date	Proportions observées			Proportions théoriques			$u = pp_0$	Indicateur de tendance
	Groupe d'âge 45-64 ans	Groupe d'âge 0-19 ans	Ensemble	Groupe d'âge 45-64 ans	Groupe d'âge 0-19 ans	Ensemble		
	p_{01}	p_{02}	p_0	p_1	p_2	p		δ
1861	0,202	0,358	0,560	0,212	0,346	0,558	0,313	3,31
1891	0,203	0,350	0,553	0,214	0,337	0,551	0,305	3,09
1911	0,208	0,336	0,544	0,220	0,326	0,546	0,297	2,89
1931	0,225	0,300	0,525	0,226	0,313	0,539	0,283	2,51
1946	0,234	0,295	0,529	0,239	0,286	0,525	0,278	2,37
1954	0,242	0,307	0,549	0,241	0,281	0,522	0,287	2,60
1962	0,229	0,331	0,560	0,241	0,275	0,516	0,289	2,67
1968	0,210	0,338	0,548	0,242	0,269	0,511	0,280	2,41
1975	0,206	0,321	0,527	0,239	0,269	0,508	0,268	2,09
1982	0,213	0,301	0,514	0,237	0,265	0,502	0,258	1,83

à 2 décimales à partir de calculs réalisés avec 4 décimales. L'évolution de l'indicateur de tendance δ est reproduite dans le graphique 1.

Signification de l'indicateur de tendance δ

La variable p_0 est égale à la probabilité de faire partie pratiquement du groupe d'âge « 0-19 ans » ou « 45-64 ans ». La variable p est la probabilité théorique du même événement, cette probabilité étant calculée à partir des observations. Le produit $u = pp_0$ représente donc la probabilité de faire partie, pratiquement et théoriquement, de ces mêmes groupes d'âge.

Les probabilités p et p_0 étant élevées (supérieures à 0,5), nous pouvons admettre que u mesure la probabilité de coïncidence des deux structures, expérimentale et théorique, de la population.

La relation trouvée indique donc que la fécondité augmente d'autant plus que la structure observée se rapproche de la structure théorique ; inversement, la fécondité diminue d'autant plus que la structure observée s'écarte de la théorique.

Conclusion

On a montré, dans cette étude, que la population générale pouvait être divisée en deux populations, pour chacune desquelles la distribution de la fécondité suit une loi normale. La fécondité de la population de moyenne la plus faible pouvait être considére comme étant celle de l'âge mûr, la fécondité de l'autre population étant celle de la jeunesse.

On en déduit un indicateur de tendance de la fécondité, fonction linéaire de la probabilité de coïncidence des structures expérimentale et théorique, de la population générale. On montre ainsi que l'augmentation de la fécondité va de pair avec le rapprochement de ces deux structures.

Annexes

1. Méthode d'ajustement

Les valeurs de la variable y sont regroupées en classes. Les effectifs des distributions D, D_1, D_2 sont notées n_j, n_{1j}, n_{2j} pour la classe j et n, n_1, n_2 pour l'ensemble.

On a :

$$n_j = n_{1j} + n_{2j}$$
$$n = n_1 + n_2$$

La proportion de l'effectif de D_1 par rapport à D est notée :

$\theta_{1j} = n_{1j}/n_j$ pour la classe j,

$\theta_1 = n_1/n$ pour l'ensemble.

Pour une distribution D à ajuster, on se donne θ_1 ainsi que la répartition des effectifs n_{1j} de D_1, pour les classes correspondant aux valeurs les plus faibles de y. Par ajustement d'une droite de Henri sur ces données, on estime les effectifs n_{1j} de D_1 pour toutes les classes. On en déduit les effectifs de D_2 par classe en effectuant les différences entre l'effectif total et l'effectif estimé de D_1. On ajuste une droite de Henri, par la méthode des moindres carrés, sur les effectifs de D_2, ainsi trouvés. On en tire une estimation des effectifs n_{2j}. On compare la distribution observée avec la somme des distributions estimées, en calculant la somme des carrés des écarts par classe.

On recommence ces opérations pour plusieurs valeurs de θ_1 et plusieurs répartitions de départ des effectifs de D_1. On conserve la structure qui donne la valeur la plus faible à la somme des carrés des écarts.

On vérifie que la somme des distributions estimées ne diffère pas significativement de la distribution observée, au niveau de confiance de 95%, en calculant la valeur du chi-deux correspondante.

Les résultats obtenus doivent être considérés comme des valeurs approchées. On constate, en effet, que la somme des carrés des écarts varie très peu pour de faibles variations de θ_1 et des effectifs supposés de D_1 au début des calculs.

2. Calcul des populations théoriques

Une table de mortalité, établie à un moment donné, s'obtient en appliquant à une génération fermée fictive les probabilités de décès par âge observées à cette date. Cette table fournit, en particulier, le nombre de survivants l_x et l'espérance de vie e_x à l'âge x.

La population théorique comprise entre l'âge x et l'âge $x + h$, calculée d'après la table, a pour effectif :

$$P(x, x+h) = l_x e_x - l_{x+h} e_{x+h}$$

En particulier, la population totale de la table est égale à :

$$P_0 = l_0 e_0$$

On en déduit la proportion de la population théorique d'âge compris entre x et $x + h$, par rapport à la population totale :

$$p(x, x+h) = P(x, x+h)/P_0$$

Références

[1] Damiani P., Massé H., Aubenque M. Etude des distributions départementales de la mortalité par cause : mise en évidence de deux populations. *Journal de la Société de Statistique de Paris*, tome 125, n° 2, 1984, 84-97.

Damiani P. Massé H., Aubenque M. Analyse de la mortalité générale par âge suivant le sexe : mise en évidence de deux types de mortalité. *Journal de la Société de statistique de Paris*, tome 125, n° 3, 1984, 158-163.

[2] Aubenque M. Indice de masculinité à la naissance. Aperçu rétrospectif et commentaires. *Journal de la Société de statistique de Paris*, tome 130, n° 2, 1989, 80-98.

[3] Mouvement de la population 1946-1947, vol 1. INSEE.

Croze M. Données de démographie régionale 1954. INSEE, 1964.

Labat J.-C. Données de démographie régionale 1962. Collection de l'INSEE, D5, 1970.

Labat J.-C., Viseur J. Données de démographie régionale 1968. Collection de l'INSEE, D23, septembre 1973.

Saboulin (de) M. Données de démographie régionale 1975. Collection de l'INSEE, D82, août 1981.

Sautory O. Données de démographie régionale 1982. Collection de l'INSEE, D 115, novembre 1986.

[4] Depoid P. Reproduction nette en Europe depuis l'origine des statistiques d'état-civil. Statistique générale de la France. Etude démographique n° 1, 1941.

[5] Liste non exhaustive :

Calot G., Deville J.-C. Nuptialité et fécondité selon le milieu socio-culturel. *Economie et statistique*, n° 27, octobre 1971.

Collomb P., Zucker E. Aspects culturels et socio-psychologiques de la fécondité française. INED. *Travaux et documents*, n° 80, 1977.

Desplanques G. Recensement général de la population de 1982. Principaux résultats de l'enquête sur les familles. Nuptialité, fécondité. Collection de l'INSEE,
D 104, 1984.

Desplanques G. Cycle de vie et milieu social. Collection de l'INSEE, D 117, février 1987.

Deville J.-C. Structure des familles. Enquête de 1962. Collection de l'INSEE, D 13-14, août 1972.

Festy P. La fécondité des pays occidentaux. INED. Travaux et documents, n° 85, 1979.

Leridon H. Aspects biométriques de la fécondité humaine. INED. *Travaux et documents*, n° 65, 1973.

Tugault Y. Fécondité et urbanisation. INED. *Travaux et documents*, n° 74, 1975.

[6] Leridon H. Fécondité et structures démographiques : une hypothèse sur l'évolution de la fécondité depuis 1940. *Population*, n° 2, 1978, 441-447.