

JOURNAL DE LA SOCIÉTÉ STATISTIQUE DE PARIS

PAUL DAMIANI

HÉLÈNE MASSÉ

MAURICE STUPFEL

Mortalité par cause et facteurs socio-démographiques

Journal de la société statistique de Paris, tome 119, n° 2 (1978), p. 131-139

http://www.numdam.org/item?id=JSFS_1978__119_2_131_0

© Société de statistique de Paris, 1978, 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/>

MORTALITÉ PAR CAUSE ET FACTEURS SOCIO-DÉMOGRAPHIQUES

Paul DAMIANI (I. N. S. E. E.), Dr Hélène MASSÉ et Dr Maurice STUPFEL
(I. N. S. E. R. M.)

On a analysé, dans cette étude, la liaison statistique existant entre les taux de mortalité par cause et certains facteurs socio-démographiques, par sexe et pour le groupe d'âge 45-64 ans. On a utilisé la méthode de la régression pas à pas. On constate la place prépondérante occupée par le facteur « alcool »; viennent ensuite les facteurs « logement surpeuplé », « cadre », population rurale ».

In this study, the statistical link between the death rates classified by causes, by sex and for the 45-64 age-group and some socio-demographic factors, has been analysed. The step wise regression method has been used. One observes the leading place occupied by the factor « Alcohol »; then come the factors « Overcrowded housing », « Executive » and « Rural population ».

INTRODUCTION

Dans cette étude, on a cherché à analyser la liaison statistique existant entre les taux de mortalité des principales causes de décès et l'ensemble des facteurs socio-démographiques pour lesquels on disposait de données. On a considéré les taux de mortalité par sexe pour le groupe d'âge 45-64 ans, rectifiés des causes de décès non spécifiées. Pour cette analyse, on a utilisé la méthode de régression pas à pas qui permet de sélectionner les facteurs dont la liaison est la plus forte.

Il convient de rappeler que de nombreux travaux ont déjà été réalisés dans ce domaine, mais ils ne concernaient en général que certains facteurs pris isolément, par exemple : l'alcool, le tabac, l'urbanisation, la catégorie socio-professionnelle, la nutrition, ... [1].

Signalons enfin qu'une étude analogue a été réalisée par H. Dubois [2]. Elle analysait, par la même méthode que celle appliquée ici, la liaison existant entre les quotients de mortalité générale, par groupe d'âge et par sexe, et l'ensemble des facteurs socio-démographiques.

DONNÉES DE BASE

Variables à expliquer : taux de mortalité par cause

Les données de base sont les statistiques départementales des causes de décès, par âge et sexe, pour la période 1968-1970 [3].

Pour éliminer l'influence de la structure par âge, on n'a considéré qu'un seul groupe d'âge. On a choisi le groupe 45-64 ans, car les décès sont suffisamment nombreux et représentatifs de la mortalité.

Les causes de décès retenues sont indiquées ci-après avec mention des numéros correspondants de la nomenclature internationale, 8^e révision, 1968 (liste abrégée et certains numéros de la liste détaillée) :

1 — Tuberculose toutes formes	B 5, B 6
2 — Cancers	B 19
3 — Diabète	B 21
4 — Maladies du cœur	B 25-B 29
5 — Lésions vasculaires cérébrales	B 30
6 — Alcoolisme et psychose alcoolique (liste détaillée)	303, 291
7 — Cirrhose du foie	B 37
8 — Accidents	BE 47, BE 48
9 — Suicide	BE 49

Correction des données

Dans la statistique des causes de décès, il existe une proportion non négligeable de décès pour lesquels la cause n'est pas spécifiée : cause non déclarée, cause mal définie (mention de manifestations purement symptomatiques le plus souvent) et décès attribués à la sénilité (rubrique B 45 de la liste abrégée). Cette proportion atteint 8,2 % des décès en 1969 et varie beaucoup d'un département à l'autre.

Pour corriger les statistiques de causes de décès des causes non spécifiées, on a utilisé une méthode inaugurée par l'I. N. S. E. E. et exposée par Ledermann [4]. On suppose que la proportion des décès de cause k dissimulés dans les décès de cause non spécifiée est constante quel que soit le département, pour le groupe d'âge et le sexe considérés. Cette proportion est alors égale au coefficient de régression changé de signe de la proportion des décès observés de cause k en fonction de la proportion des décès de cause non spécifiée, calculé à partir des données départementales. On rectifie les décès observés de cause k en y ajoutant la proportion ainsi calculée des décès de cause non spécifiée attribuables à la cause k .

On calcule des taux de mortalité rectifiés par cause en divisant les nombres corrigés de décès par la population correspondante. On dispose donc, pour chaque sexe et par département, des taux rectifiés de mortalité du groupe d'âge 45-64 ans, pour chacune des 9 causes définies précédemment. Ces taux, calculés pour 100 000 habitants, sont notés y_k ($k = 1, 2, \dots, 9$).

Variables explicatives : facteurs socio-démographiques

On a retenu les principales variables socio-démographiques pour lesquelles on disposait de données départementales [5]. Ce sont les mêmes quel que soit le sexe de la population considérée. Ces variables notés x_j , ($j = 1, 2, \dots, 17$), sont les suivantes :

Urbanisation

x_1 = proportion pour 100 de la population rurale (population des communes de moins de 2 000 habitants) en 1968.

Revenus

x_2 = revenu en milliers de francs, par cote, en 1968.

Logement

x_3 = proportion pour 100 des logements ayant certains éléments de confort en 1968,

x_4 = proportion pour 100 des logements ayant un surpeuplement accentué en 1968.

Instruction

x_5 = proportion pour 100 des personnes de 15 ans ou plus possédant un diplôme supérieur au baccalauréat en 1968.

Catégorie socio-professionnelle

Proportion pour 100 des ménages ordinaires, suivant la catégorie socio-professionnelle du chef de ménage, en 1968 :

x_6 = agriculteur exploitant,

x_7 = salarié agricole,

x_8 = patron de l'industrie et du commerce,

x_9 = profession libérale et cadre supérieur,

x_{10} = cadre moyen,

x_{11} = employé,

x_{12} = ouvrier.

Étrangers

x_{13} = proportion pour 100 des étrangers en 1968.

Équipement sanitaire

x_{14} = nombre de médecins pour 100 000 habitants en 1970,

x_{15} = nombre de lits du secteur hôpital, du secteur public et du secteur privé, pour 1 000 habitants en 1971.

Consommation

x_{16} = comme la consommation d'alcool par département est inconnue, on a pris, comme indicateur de cette consommation, une moyenne pondérée des taux de mortalité par alcoolisme et par cirrhose du foie suivant le sexe, du groupe d'âge 45-64 ans, pour la période 1968-1970, ces taux étant rectifiés pour tenir compte des causes non spécifiées. Les coefficients de pondération sont proportionnels aux décès dus à l'alcoolisme et à 70 pour cent des décès dus à la cirrhose du foie.

x_{17} = vente de tabac en hg par habitant en 1967.



MÉTHODE UTILISÉE

On a utilisé la méthode de régression pas à pas. Les calculs ont été réalisés sur ordinateur.

Régression linéaire [6]

On suppose qu'il existe un modèle de régression linéaire entre le taux de mortalité de la cause k , pour un sexe donné, et l'ensemble des variables x_j . Ce modèle s'écrit :

$$y_k = a_k + b_{k1} x_1 + \dots + b_{kj} x_j + \dots + e_k$$

où : a_k est une constante,

b_{kj} est le coefficient de régression relatif au facteur x_j ,

e_k est le résidu.

On suppose que, pour l'ensemble des départements, les résidus sont d'espérance mathématique nulle, sont sans corrélation entre eux et ont même variance.

Les coefficients a_k , b_{kj} sont déterminés par la méthode des moindres carrés en écrivant que la somme des carrés des résidus est minimum.

On peut décomposer la variance de y_k en deux parties :

- une partie expliquée par le modèle,
- la variance résiduelle.

On appelle R^2 la proportion de la variance de y qui est expliquée par le modèle. C'est le carré du coefficient de corrélation multiple de y en fonction des variables x_j . La valeur de R^2 est inférieure ou égale à 1. Le modèle est d'autant meilleur que la valeur de R^2 est plus grande.

Parmi l'ensemble des variables x_j , certaines peuvent être éliminées car elles n'améliorent pas le modèle de manière significative. Pour déterminer les variables à retenir, on utilise la méthode de la régression pas à pas.

Régression pas à pas

On calcule successivement le modèle de régression avec une première variable explicative, puis avec deux, trois, etc... On s'arrête à p variables quand on juge que l'apport d'une $(p + 1)^e$ variable n'améliore pas le modèle.

Le choix des variables à prendre dans le modèle se fait de la façon suivante :

- la première variable choisie est celle dont le coefficient de corrélation avec y est maximum;
- d'une manière générale, si $(p - 1)$ variables ont déjà été introduites dans le modèle, la p^e variable choisie est, parmi les variables restantes, celle dont l'introduction dans le modèle apporte un accroissement maximum du coefficient de corrélation multiple.

Pratiquement, on a adopté la règle générale suivante : on n'a pas retenu une $(p + 1)^e$ variable si l'accroissement relatif de la part de la variance expliquée résultant de son introduction dans le modèle était inférieur à 10 %. Pour l'alcoolisme et la cirrhose, on a retenu une règle moins restrictive par suite de la prépondérance du facteur « alcool ».

RÉSULTATS

Le tableau 1 fournit les coefficients de corrélation entre les taux de mortalité par cause, suivant le sexe, et les facteurs socio-démographiques. Ils donnent une première indication sur la liaison statistique apparente existant entre ces variables.

Les tableaux 2 donnent, suivant le sexe, par cause de décès, les modèles de régression retenus. On a indiqué, pour chaque cause, le modèle correspondant à la prise en compte de chacun des facteurs successifs, avec l'indication de la part de la variance du taux de mortalité expliquée par le modèle correspondant.

Si on compare les résultats par sexe, on observe que, pour le sexe masculin, en général, l'ajustement est meilleur et nécessite moins de facteurs explicatifs.

TABLEAU 1

*Coefficients de corrélation entre les taux de mortalité par cause,
suivant le sexe, et les facteurs socio-démographiques*

Facteurs socio-démographiques	Causes de décès								
	Tuber- culose	Cancers	Diabète	Maladies du cœur	Lésions vasculaires cérébrales	Alcoolisme	Cirrhose du foie	Accidents	Suicide
<i>Sexe masculin</i>									
x_1 = population rurale	0,110	-0,316	0,122	-0,085	0,356	0,290	-0,066	0,153	0,418
x_2 = revenu	-0,058	-0,062	0,002	-0,069	-0,128	-0,178	-0,127	0,006	-0,204
x_3 = logement confortable	-0,287	0,046	-0,079	0,026	-0,419	-0,288	-0,151	-0,091	-0,486
x_4 = logement surpeuplé	0,159	0,124	-0,305	-0,273	0,431	0,402	0,320	0,269	0,396
x_5 = diplômés	-0,218	0,190	-0,098	-0,106	-0,357	-0,316	-0,114	-0,144	-0,441
x_6 = agriculteur exploitant	0,118	-0,327	0,118	-0,121	0,380	0,175	-0,150	-0,018	0,252
x_7 = salarié agricole	-0,152	-0,230	0,034	-0,386	-0,062	-0,144	-0,316	-0,147	0,149
x_8 = patron	-0,137	-0,402	0,055	-0,308	0,158	-0,034	-0,319	-0,035	0,065
x_9 = profession libérale	-0,209	0,193	-0,149	-0,097	-0,362	-0,248	0,001	-0,084	-0,367
x_{10} = cadre moyen	-0,218	0,237	-0,196	-0,045	-0,410	-0,193	0,060	-0,015	-0,335
x_{11} = employé	-0,250	0,275	-0,147	-0,017	-0,465	-0,301	-0,001	-0,152	-0,413
x_{12} = ouvrier	0,085	0,422	-0,049	0,275	0,108	0,131	0,366	0,339	0,115
x_{13} = étrangers	-0,241	-0,024	0,064	0,132	-0,459	-0,397	-0,253	-0,325	-0,608
x_{14} = équipement médical	-0,152	0,059	-0,011	-0,057	-0,216	-0,370	-0,181	-0,210	-0,410
x_{15} = équipement hospitalier	-0,064	0,101	0,039	0,175	-0,170	-0,186	-0,057	-0,108	-0,289
x_{16} = alcool	0,568	0,634	0,037	0,420	0,539	0,770	0,972	0,622	0,533
x_{17} = tabac	0,098	0,183	0,027	0,220	-0,334	-0,102	-0,030	-0,088	-0,333
<i>Sexe féminin</i>									
x_1 = population rurale	-0,225	-0,282	-0,043	0,077	0,401	0,057	-0,196	-0,061	0,234
x_2 = revenu	0,104	0,069	-0,062	-0,164	-0,167	-0,033	0,028	0,145	-0,097
x_3 = logement confortable	0,174	-0,011	-0,001	-0,145	-0,401	-0,174	-0,084	0,016	-0,346
x_4 = logement surpeuplé	0,256	-0,145	-0,332	-0,121	0,111	0,371	0,284	0,059	0,219
x_5 = diplômés	0,114	0,231	-0,177	-0,283	-0,434	-0,137	0,043	0,015	-0,247
x_6 = agriculteur exploitant	-0,254	-0,193	-0,137	0,038	0,391	-0,045	-0,282	-0,225	0,017
x_7 = salarié agricole	-0,045	-0,106	-0,001	-0,239	-0,031	-0,162	-0,200	-0,180	-0,092
x_8 = patron	0,006	-0,240	-0,260	-0,227	0,012	-0,265	-0,426	-0,146	-0,052
x_9 = profession libérale	0,130	0,279	-0,182	-0,274	-0,455	-0,017	0,160	0,083	-0,167
x_{10} = cadre moyen	0,123	0,299	-0,172	-0,212	-0,431	0,102	0,253	0,160	-0,111
x_{11} = employé	0,121	0,274	-0,071	-0,203	-0,491	-0,010	0,194	0,159	-0,148
x_{12} = ouvrier	0,087	0,114	0,085	0,213	0,083	0,303	0,442	0,312	0,188
x_{13} = étrangers	0,109	-0,027	0,144	-0,016	-0,304	-0,287	-0,136	-0,108	-0,445
x_{14} = équipement médical	0,125	0,155	-0,046	-0,234	-0,351	-0,273	-0,031	-0,013	-0,277
x_{15} = équipement hospitalier	0,039	0,083	0,053	-0,019	-0,257	-0,188	-0,037	-0,020	-0,206
x_{16} = alcool	0,409	0,164	0,182	0,365	0,361	0,663	0,705	0,315	0,280
x_{17} = tabac	0,170	0,058	0,110	0,046	-0,309	-0,116	-0,007	0,004	-0,210

TABLEAU 2.1

Modèles de régression suivant la cause de décès

Sexe masculin

Variables explicatives	Modèles (1)	R ² proportion de la variance expliquée
<i>1 — Tuberculose</i>		
x_{16} = alcool	$y_1 = 8,08949 + 0,22890 x_{16}$	0,323
x_{10} = cadre	$y_1 = 14,16228 + 0,23406 x_{16} - 0,98883 x_{10}$	0,384
x_{17} = tabac	$y_1 = -4,28878 + 0,24168 x_{16} - 1,45285 x_{10} + 1,55831 x_{17}$ (6,56856) (0,03155) (0,34330) (0,48331)	0,447
<i>2 — Cancers</i>		
x_{16} = alcool	$y_2 = 271,35465 + 1,46560 x_{16}$	0,401
x_1 = population rurale	$y_2 = 312,34701 + 1,45027 x_{16} - 0,89630 x_1$ (17,40728) (0,17183) (0,22055)	0,493
<i>3 — Diabète</i>		
x_4 = logement surpeuplé	$y_3 = 20,25404 - 0,54066 x_4$	0,093
x_{10} = cadre	$y_3 = 23,58307 - 0,58868 x_4 - 0,44369 x_{10}$	0,147
x_{16} = alcool	$y_3 = 21,95426 - 0,72651 x_4 - 0,48062 x_{10} + 0,03881 x_{16}$ (2,20598) (0,18299) (0,18196) (0,01966)	0,182
<i>4 — Maladies du cœur</i>		
x_{16} = alcool	$y_4 = 183,35618 + 0,79362 x_{16}$	0,177
x_4 = logement surpeuplé	$y_4 = 231,86510 + 1,14392 x_{16} - 8,70715 x_4$	0,391
x_7 = salarié agricole	$y_4 = 251,38289 + 0,99768 x_{16} - 8,12457 x_4 - 4,37337 x_7$	0,435
x_{10} = cadre	$y_4 = 296,76856 + 0,97783 x_{16} - 8,56343 x_4 - 6,64864 x_7 - 5,06367 x_{10}$ (20,99151) (0,16115) (1,43257) (1,70361) (1,55511)	0,495
<i>5 — Lésions vasculaires cérébrales</i>		
x_{16} = alcool	$y_5 = 70,93411 + 0,59319 x_{16}$	0,291
x_6 = agriculteur exploitant	$y_5 = 48,10306 + 0,65454 x_{16} + 1,23323 x_6$ (7,71774) (0,08222) (0,20324)	0,494
<i>6 — Alcoolisme</i>		
x_{16} = alcool	$y_6 = -8,21143 + 0,59348 x_{16}$	0,592
x_1 = population rurale	$y_6 = -22,12363 + 0,59869 x_{16} + 0,30419 x_1$ (4,56122) (0,04502) (0,05779)	0,687
<i>7 — Cirrhose du foie</i>		
x_{16} = alcool	$y_7 = 6,76265 + 1,49015 x_{16}$ (3,11333) (0,03770)	0,944
<i>8 — Accidents</i>		
x_{16} = alcool	$y_8 = 72,14080 + 0,74726 x_{16}$ (8,06403) (0,09765)	0,386
<i>9 — Suicide</i>		
x_{13} = étrangers	$y_9 = 67,63849 - 3,39509 x_{13}$	0,369
x_{16} = alcool	$y_9 = 42,28257 - 2,75602 x_{13} + 0,28346 x_{16}$ (5,48827) (0,42831) (0,05614)	0,506

1. Sous les coefficients du dernier modèle retenu par cause, figurent, entre parenthèses, les erreurs-type de ces coefficients.

Si on met à part la mortalité par alcoolisme et par cirrhose du foie, on constate que la part de la variance expliquée atteint les valeurs suivantes :

- pour le sexe masculin, 18 % pour le diabète, entre 40 et 50 % pour les autres causes de décès,
- pour le sexe féminin, 17 % pour les accidents, 23 % pour les cancers, entre 30 et 40 % pour les autres causes de décès.

Pour ce qui concerne les facteurs socio-démographiques retenus, on peut faire les observations suivantes :

- pour le sexe masculin, 8 facteurs seulement sont représentés : le facteur « alcool » est présent dans toutes les causes de décès, le facteur « cadre » est présent 3 fois, les facteurs « population rurale » et « logement surpeuplé » 2 fois;

- pour le sexe féminin, 15 facteurs sont représentés : le facteur « alcool » intervient dans 7 causes sur 9, le facteur « logement surpeuplé » dans 4 causes, les facteurs « logement confortable » et « cadre » dans 3 causes, les facteurs « population rurale », « agriculteur », « patron », « ouvrier » dans 2 causes.

TABLEAU 2.2

Modèles de régression suivant la cause de décès

Sexe féminin

Variables explicatives	Modèles (1)	proportion de la variance expliquée
	1 — Tuberculose	
x_{16} = alcool	$y_1 = 3,21191 + 0,05579 x_{16}$	0,167
x_3 = logement confortable . . .	$y_1 = -0,26335 + 0,06243 x_{16} + 0,07380 x_3$	0,232
x_6 = patron	$y_1 = -5,78539 + 0,07309 x_{16} + 0,08740 x_3 + 0,43604 x_6$	0,274
x_8 = agriculteur exploitant . . .	$y_1 = -4,52177 + 0,06934 x_{16} + 0,03579 x_3 + 0,72773 x_6 - 0,10887 x_8$ (2,85543) (0,01306) (0,03434) (0,22600) (0,04753)	0,314
	2 — Cancers	
x_{10} = cadre	$y_2 = 200,35905 + 3,19147 x_{10}$	0,089
x_3 = logement confortable . . .	$y_2 = 215,70040 + 5,50746 x_{10} - 0,76384 x_3$	0,158
x_4 = logement surpeuplé . . .	$y_2 = 243,01076 + 5,94753 x_{10} - 1,00488 x_3 - 2,34680 x_4$	0,206
x_{15} = équipement hospitalier . .	$y_2 = 231,67644 + 6,35668 x_{10} - 1,21149 x_3 - 2,36075 x_4 + 2,50457 x_{15}$ (15,75790) (1,31116) (0,30927) (0,99068) (1,38694)	0,234
	3 — Diabète	
x_4 = logement surpeuplé . . .	$y_3 = 20,07472 - 0,69539 x_4$	0,110
x_{16} = alcool	$y_3 = 16,38528 - 0,65447 x_4 + 0,07514 x_{16}$	0,205
x_{10} = cadre	$y_3 = 20,25109 - 1,03588 x_4 + 0,08131 x_{16} - 0,55562 x_{10}$	0,265
x_1 = population rurale	$y_3 = 31,25152 - 0,92587 x_4 + 0,07904 x_{16} - 1,42129 x_{10} - 0,13768 x_1$	0,333
x_6 = patron	$y_3 = 44,26999 - 0,68882 x_4 + 0,04840 x_{16} - 1,80267 x_{10} - 0,14242 x_1 - 1,06310 x_6$ (6,23829) (0,21035) (0,02310) (0,36046) (0,04389) (0,37793)	0,387
	4 — Maladies du cœur	
x_{16} = alcool	$y_4 = 67,26709 + 0,27368 x_{16}$	0,133
x_4 = logement surpeuplé . . .	$y_4 = 75,80894 + 0,35702 x_{16} - 2,07171 x_4$	0,210
x_9 = profession libérale	$y_4 = 89,87005 + 0,36006 x_{16} - 2,26249 x_4 - 2,07134 x_9$	0,295
x_7 = salarié agricole	$y_4 = 102,02277 + 0,29500 x_{16} - 2,05737 x_4 - 2,69806 x_9 - 1,97317 x_7$ (8,52377) (0,07329) (0,64524) (0,65156) (0,75561)	0,345
	5 — Lésions vasculaires cérébrales	
x_{11} = employé	$y_5 = 88,35061 - 3,39446 x_{11}$	0,241
x_{16} = alcool	$y_5 = 74,19894 - 3,32791 x_{11} + 0,17300 x_{16}$ (5,13055) (0,57604) (0,04152)	0,362
	6 — Alcoolisme	
x_{16} = alcool	$y_6 = -3,32389 + 0,13739 x_{16}$	0,440
x_4 = logement surpeuplé . . .	$y_6 = -4,85780 + 0,12632 x_{16} + 0,27533 x_4$	0,458
x_{14} = équipement médical . . .	$y_6 = -2,74971 + 0,11984 x_{16} + 0,28143 x_4 - 0,01456 x_{14}$ (1,96040) (0,01733) (0,15689) (0,00815)	0,476
	7 — Cirrhose du foie	
x_{16} = alcool	$y_7 = -4,05919 + 0,61251 x_{16}$	0,497
x_{10} = cadre	$y_7 = -15,54072 + 0,60276 x_{16} + 1,86953 x_{10}$	0,544
x_3 = logement confortable . . .	$y_7 = -6,89533 + 0,57019 x_{16} + 2,81779 x_{10} - 0,30725 x_3$	0,560
x_{12} = ouvrier	$y_7 = -9,15334 + 0,51914 x_{16} + 2,45104 x_{10} - 0,83713 x_3 + 0,30539 x_{12}$ (7,80453) (0,08747) (0,80743) (0,16853) (0,15694)	0,578
	8 — Accidents	
x_{16} = alcool	$y_8 = 24,05199 + 0,12720 x_{16}$	0,099
x_{13} = ouvrier	$y_8 = 20,05714 + 0,09326 x_{16} + 0,20617 x_{13}$	0,143
x_9 = revenus	$y_8 = 15,07512 + 0,10467 x_{16} + 0,18629 x_{13} + 0,37619 x_9$ (4,71504) (0,04210) (0,09486) (0,22356)	0,169
	9 — Suicide	
x_{13} = étrangers	$y_9 = 20,23752 - 0,97095 x_{13}$	0,198
x_6 = agriculteur exploitant . . .	$y_9 = 26,03841 - 1,43006 x_{13} - 0,25652 x_6$	0,235
x_1 = population rurale	$y_9 = 21,09535 - 1,21153 x_{13} - 0,47061 x_6 + 0,15952 x_1$ (2,69183) (0,24277) (0,10913) (0,05973)	0,337

1. Sous les coefficients du dernier modèle retenu par cause, figurent, entre parenthèses, les erreurs-type de ces coefficients.

Comparaison avec l'étude de H. Dubois

L'étude de H. Dubois déjà citée [2] appliquait la même méthode d'analyse aux quotients de mortalité générale, par groupe d'âge et par sexe, les facteurs socio-démographiques retenus étant sensiblement les mêmes que ceux utilisés ici. Les résultats obtenus, dans cette étude, pour le groupe d'âge 45-64 ans, sont les suivants :

- pour le sexe masculin, on explique 57 % de la variance du quotient de mortalité, avec un modèle comprenant 3 facteurs (« alcool », « population rurale », « équipement médical »);
- pour le sexe féminin, on explique 35 % de la variance avec un modèle à 4 facteurs (« alcool », « population rurale », « équipement médical », « tabac »).

Ces résultats sont en concordance avec ceux de la présente étude.

CONCLUSION

La méthode d'analyse statistique adoptée dans cette étude permet de faire apparaître les différents facteurs socio-démographiques qui interviennent pour expliquer la dispersion, autour de la moyenne, des taux départementaux de mortalité par cause. Ces facteurs apparaissent dans le modèle dans l'ordre décroissant de la part qu'ils prennent à cette explication.

Si on considère les résultats obtenus pour les deux sexes, on observe la place prépondérante occupée par le facteur « alcool ». Viennent ensuite les facteurs « logement surpeuplé », « cadre », « population rurale ».

RÉFÉRENCES

[1] On peut citer notamment :

AUBENQUE M., DAMIANI P., DERUFFE L. — Essai d'appréciation de l'incidence de la grippe sur les fluctuations de la mortalité. *Études et conjoncture*. I. N. S. E. E., n° 9, 1965, 121-127.

DAMIANI P. — La mortalité départementale attribuée au cancer broncho-pulmonaire et la consommation de tabac. *Études et conjoncture*. I. N. S. E. E., n° 12, 1965, 137-142.

MASSÉ H. — Liaison statistique entre la mortalité par cancer, la mortalité par maladies cardiaques et certains facteurs socio-démographiques. *Bulletin de l'I. N. S. E. R. M.*, t. 25, n° 2, 1970, 295-306.

MASSÉ H. — L'alcoolisme, facteur de mortalité : son incidence sur les principales causes de décès. *La nouvelle presse médicale*, 1^{er} juillet 1972, 1, n° 27, 1857-1860.

MASSÉ L., DAMIANI P., MASSÉ H. — Le tabac et l'urbanisation, facteurs de mortalité. *Bulletin de l'Académie nationale de médecine*, t. 158, n° 5, séance du 21 mai 1974, 354-361.

MASSÉ H. — Nutrition et mortalité : influence des constituants alimentaires sur les principales causes de décès. *Journal de la Société de statistique de Paris*, t. 117, n° 2, 1976, 154-159.

MASSÉ H. — Liaison entre la mortalité par cause et la catégorie socio-professionnelle. *Journal de la Société de statistique de Paris*, t. 118, n° 2, 1977, 165-169.

- [2] DUBOIS H. — Influence de certains facteurs sur la mortalité. *Bulletin de l'Institut des actuaires français* (sous presse).
- [3] Statistiques des causes médicales de décès. Volumes annuels. I. N. S. E. R. M.
- [4] LEDERMANN S. — La répartition des décès de cause indéterminée. *Revue de l'Institut international de statistique*. I-III, 47-57, 1956.
- [5] Annuaire statistique de la France. I. N. S. E. E.
Résultats du recensement de la population de 1968. I. N. S. E. E.
CROZE M. — Tableaux démographiques et sociaux. I. N. S. E. E.-I. N. E. D. 1976.
La consommation du tabac en France. S. E. I. T. A.
- [6] DAMIANI P. — Méthodes d'analyse statistique. I. N. S. E. E. 1976.