

PAUL DAMIANI

HÉLÈNE MASSÉ

MAURICE AUBENQUE

**Étude des distributions départementales de la mortalité par cause : mise en évidence de deux populations**

*Journal de la société statistique de Paris*, tome 125, n° 2 (1984), p. 84-97

[http://www.numdam.org/item?id=JSFS\\_1984\\_\\_125\\_2\\_84\\_0](http://www.numdam.org/item?id=JSFS_1984__125_2_84_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/>

## II

## ARTICLES

**ÉTUDE DES DISTRIBUTIONS DÉPARTEMENTALES DE LA MORTALITÉ  
PAR CAUSE : MISE EN ÉVIDENCE DE DEUX POPULATIONS**

Paul DAMIANI (I.N.S.E.E.), Hélène MASSÉ (I.N.S.E.R.M.), Maurice AUBENQUE (I.N.S.E.E.) <sup>(1)</sup>

*Les auteurs ont étudié les distributions départementales des proportions de décès par cause par rapport à l'ensemble des décès, par sexe, pour le groupe d'âge 45-64 ans. Ils ont constaté que ces distributions étaient la somme de deux distributions log-normales. Ils en ont déduit l'existence de deux populations suivant que la nature de la mortalité pour une cause donnée était d'origine endogène ou exogène.*

*In this paper, the authors analyse the distributions by French « département » of proportions of deaths by cause, according to sex and 45-64 years age-group. They find that they are sum of two log-normal distributions. They infer that population can be decomposed into two sets according the nature of mortality, endogenous or exogenous.*

## INTRODUCTION

Dans cette étude, on a essayé d'ajuster des lois de probabilité connues sur les distributions départementales des proportions de décès par cause par rapport au total des décès. Cette analyse a été réalisée par sexe, pour le groupe d'âge 45-65 ans.

Dans un travail précédent [1], P. Damiani, H. Massé et M. Aubenque avaient réalisé un tel ajustement sur la proportion de décès par alcoolisme et psychose alcoolique. Ils avaient constaté que la distribution correspondante des valeurs départementales était la somme de deux distributions log-normales. Ils en avaient déduit l'existence de deux populations différant suivant l'origine endogène ou exogène de la propension à consommer de l'alcool.

On montrera, dans cet article, qu'on obtient des résultats analogues pour toutes les causes de décès.

## DONNÉES DE BASE

Les données de base sont les statistiques départementales des causes de décès pour la période 1968-1970 [2]. On a utilisé les décès, par sexe, pour le groupe d'âge 45-64 ans. On a retenu neuf causes de décès dont la liste figure dans le tableau 1, avec les numéros correspondants de la classification inter-

---

1. Institut national de la statistique et des études économiques (I.N.S.E.E.), 18, boulevard A.-Pinard, 75675 Paris Cedex 14.  
Institut national de la santé et de la recherche médicale (I.N.S.E.R.M.), 101, rue de Tolbiac, 75654 Paris Cedex 13.

nationale, révision 1965. On a corrigé les statistiques pour tenir compte des décès de cause non spécifiée, suivant une méthode inaugurée par Breil à l'I.N.S.E.E. [3] et exposée par Ledermann [4]. A partir des données ainsi rectifiées, on calcule la proportion des décès de chaque cause par rapport à l'ensemble des décès par département, suivant le sexe, pour le groupe d'âge considéré. Cette proportion de décès est appelée *mortalité relative* de la maladie considérée. On est amené à corriger les valeurs de certains départements jugées aberrantes, par comparaison avec les valeurs d'autres départements de même composition urbaine et proches géographiquement.

On admet que, pour chaque département, par sexe et par groupe d'âge, le taux de mortalité pour une cause donnée, rapporté au nombre de malades correspondant, est proportionnel au taux de mortalité générale. On en déduit que la mortalité relative d'une maladie donnée est proportionnelle au *taux de prévalence* correspondant, c'est-à-dire au nombre de malades par habitant pour cette affection (voir annexe 1).

Si on appelle  $x_{kj}$  et  $n_{kj}$  la mortalité relative et le taux de prévalence de la maladie  $k$  pour le département  $j$ , on a, par sexe, la relation :

$$x_{kj} = \gamma_k n_{kj} \quad (1)$$

où  $\gamma_k$  est un coefficient dépendant de la maladie et du sexe et indépendant du département.

La valeur de  $\gamma_k$  se calcule en appliquant la relation (1) à la France entière et en utilisant les résultats d'une étude de P. Damiani, H. Massé et M. Aubenque, sur l'évaluation de la morbidité à partir de la mortalité [5].

#### AJUSTEMENT DES DISTRIBUTIONS DÉPARTEMENTALES

On dispose, pour 95 départements, des valeurs de la mortalité relative pour 9 causes de décès, suivant le sexe et pour le groupe d'âge 45-64 ans.

On essaie, pour chaque cause et par sexe, d'ajuster une loi sur la distribution correspondante. On constate que l'ajustement d'une loi simple de probabilité n'est pas satisfaisant. Comme dans le cas de la mortalité relative par alcoolisme étudié précédemment [1], on obtient un meilleur résultat en considérant que chacune de ces distributions est la somme de deux distributions log-normales. Plus précisément, si on prend comme nouvelle variable, le logarithme de la variable précédente, les distributions observées sont la somme de deux distributions normales.

L'ajustement se fait par approximations successives (voir annexe 2).

Pour une cause donnée et par sexe, on calcule, pour chacune des distributions composantes, l'importance relative des effectifs par rapport à l'ensemble, la répartition des fréquences suivant les valeurs de la variable ainsi que la moyenne et l'écart type. Les distributions composantes sont appelées  $D_1$  et  $D_2$ ;  $D_1$  étant la distribution de moyenne la plus faible,  $D_2$  celle de moyenne la plus élevée (voir annexe 3).

La relation (1) permet de déterminer les caractéristiques de ces distributions par rapport au taux de prévalence.

#### APPLICATIONS

L'hypothèse faite précédemment montre que les distributions départementales du taux de prévalence sont également la somme de deux distributions log-normales.

Les valeurs départementales du logarithme du taux de prévalence sont approximativement les

moyennes des logarithmes de cette variable. On a trouvé que la distribution de ces valeurs était la somme de deux lois normales. On peut en déduire que la distribution des logarithmes du taux de prévalence calculé sur la population générale est également la somme de deux lois normales de même importance relative et de mêmes moyennes que celles des distributions départementales.

Autrement dit, on peut admettre que la population générale  $P$  se décompose en deux populations  $P_1$  et  $P_2$  pour chacune desquelles la distribution du taux de prévalence de la maladie  $k$  suit une loi log-normale. La répartition des effectifs de ces populations par rapport à la population totale ainsi que les moyennes de ces distributions sont celles trouvées précédemment pour  $D_1$  et  $D_2$ .

L'étude citée en [5] fournit, par ailleurs, des estimations, par maladie, des taux de prévalence et des taux de mortalité rapportés au nombre de malades correspondants. Les résultats obtenus dans la présente étude permettent d'évaluer ces taux pour chacune des populations composantes (voir annexe 4).

## RÉSULTATS

Le tableau 1 donne, suivant la maladie, par sexe et pour le groupe d'âge 45-64 ans, l'importance relative, le taux de prévalence et le taux de mortalité rapporté au nombre de malades, pour les populations  $P_1$  et  $P_2$ .

Les tableaux 2 fournissent, suivant la maladie, par sexe et pour le groupe d'âge 45-64 ans, la distribution des fréquences ajustées des proportions départementales de décès par cause, pour  $D_1$ ,  $D_2$  et pour l'ensemble.

Des graphiques illustrent ces résultats. On notera que les échelles employées pour représenter les distributions des deux sexes d'une même maladie sont les mêmes; par contre, les échelles diffèrent d'une maladie à l'autre.

Si on classe les maladies, par ordre croissant d'importance de la population  $P_2$ , on obtient les résultats suivants, pour le sexe masculin :

- 40 % maladies du cœur,
- 50 % tuberculose, cancers, maladies cérébro-vasculaires, accidents,
- 55 % cirrhose du foie,
- 60 % suicide,
- 70 % diabète, alcoolisme.

L'importance relative de  $P_2$ , pour le sexe féminin, est égale à celle du sexe masculin, pour les maladies cérébro-vasculaires (50 %), le diabète (70 %); elle est inférieure pour l'alcoolisme (M = 70 %, F = 30 %), la tuberculose (M = 50 %, F = 45 %), la cirrhose du foie (M = 55 %, F = 50 %), le suicide (M = 60 %, F = 50 %); elle est supérieure pour les cancers (M = 50 %, F = 60 %), les maladies du cœur (M = 40 %, F = 55 %), les accidents (M = 50 %, F = 60 %).

## INTERPRÉTATION DES RÉSULTATS

Pour interpréter les résultats obtenus, on propose une hypothèse analogue à celle avancée dans le cas de l'étude de la mortalité par alcoolisme [1]. La mortalité de la population  $P_1$  a une origine exogène, due à des circonstances extérieures, à l'environnement, au milieu... Celle de la population  $P_2$  est d'ori-

gine endogène, relativement indépendante d'une atteinte extérieure à l'organisme lui-même, telle qu'une maladie acquise au cours de l'existence, d'un traumatisme, d'accidents, etc.

D'une manière plus précise, la discrimination entre les deux populations  $P_1$  et  $P_2$  résulte d'une analyse statistique qui met en évidence la prédominance de l'un des deux facteurs endogène ou exogène.

Par exemple, pour les maladies du cœur du sexe masculin, l'étude montre que 40 % des cardiopathies fatales se développent sur un terrain prédisposé (y compris, bien entendu, les cardiopathies effectivement congénitales, telles que les malformations cardiaques); les autres maladies du cœur résultant, au premier chef, d'une affection acquise (atteinte rhumatismale, infectieuse...), l'athérosclérose participant bien probablement des deux processus.

TABLEAU 1

*Importance relative, taux de prévalence et taux de mortalité des populations  $P_1$  et  $P_2$ , par maladie, suivant le sexe, pour le groupe d'âge 45-64 ans*

Affections	Numéros de la classification internationale (révision 1965) (1)	Sexe	Populations	Importance relative en %	Taux de prévalence pour 1 000	Taux de mortalité rapporté aux malades pour 1 000
Tuberculose toutes formes . . .	B 5, B 6	M	$P_1$	50	12,1	12,9
		F	$P_2$	50	12,1	16,9
Cancers toutes formes . . . . .	B 19	M	$P_1$	55	6,9	4,7
			$P_2$	45	5,6	6,7
		F	$P_1$	50	101,1	18,3
			$P_2$	50	101,1	19,9
Maladies cérébro-vasculaires . .	B 30	M	$P_1$	40	54,0	20,3
			$P_2$	60	81,0	14,6
		F	$P_1$	50	35,7	8,7
			$P_2$	50	35,7	9,7
Maladies du cœur . . . . .	B 25 - B 29	M	$P_1$	60	86,7	13,8
			$P_2$	40	57,8	21,3
		F	$P_1$	45	61,4	6,6
			$P_2$	55	75,0	5,9
Diabète . . . . .	B 21	M	$P_1$	30	3,3	25,1
			$P_2$	70	7,7	12,7
		F	$P_1$	30	8,1	9,0
			$P_2$	70	18,8	5,7
Alcoolisme . . . . .	291, 303	M	$P_1$	30	15,6	13,5
			$P_2$	70	36,5	7,6
		F	$P_1$	70	11,9	4,3
			$P_2$	30	5,1	11,7
Cirrhose du foie . . . . .	B 37	M	$P_1$	45	55,1	12,6
			$P_2$	55	67,3	11,3
		F	$P_1$	50	45,2	5,3
			$P_2$	50	45,2	6,7
Suicide . . . . .	BE 49	M	$P_1$	40	14,6	15,6
			$P_2$	60	21,8	15,3
		F	$P_1$	50	7,0	12,0
			$P_2$	50	7,0	14,0
Accidents . . . . .	BE 47, BE 48	M	$P_1$	50	40,1	15,7
			$P_2$	50	40,1	17,5
		F	$P_1$	40	11,7	12,9
			$P_2$	60	17,5	11,6

1. Liste B et certains numéros de la liste détaillée.

TABLEAUX 2

*Distribution des fréquences ajustées  
des proportions départementales de décès par cause,  
suivant le sexe, pour le groupe d'âge 45-64 ans.*

## 1. Tuberculose

Proportion de décès par tuberculose (en %)	Fréquences ajustées		
	$D_1$	$D_2$	Ensemble
<i>Sexe masculin</i>			
Moins de 1,2	0,021	0,009	0,030
1,2 - 1,3	0,075	0,023	0,098
1,4 - 1,5	0,127	0,044	0,171
1,6 - 1,7	0,126	0,063	0,189
1,8 - 1,9	0,083	0,072	0,155
2,0 - 2,1	0,042	0,070	0,112
2,2 - 2,3	0,017	0,061	0,078
2,4 - 2,5	0,006	0,049	0,055
2,6 et plus	0,003	0,109	0,112
Ensemble	0,500	0,500	1,000
Moyenne	1,67	2,19	1,89
Écart type	0,30	0,60	0,44
<i>Sexe féminin</i>			
Moins de 0,6	0,041	0,020	0,061
0,6 - 0,7	0,095	0,053	0,148
0,8 - 0,9	0,118	0,077	0,195
1,0 - 1,1	0,103	0,079	0,182
1,2 - 1,3	0,075	0,067	0,142
1,4 - 1,5	0,049	0,050	0,099
1,6 - 1,7	0,029	0,036	0,065
1,8 - 1,9	0,017	0,024	0,041
2,0 et plus	0,023	0,044	0,067
Ensemble	0,550	0,450	1,000
Moyenne	1,12	1,29	1,19
Écart type	0,44	0,54	0,46

## 2. Cancers

Proportion de décès par cancers (en %)	Fréquences ajustées		
	$D_1$	$D_2$	Ensemble
<i>Sexe masculin</i>			
Moins de 23,0	0,003	0,002	0,005
23,0 - 24,9	0,046	0,016	0,062
25,0 - 26,9	0,162	0,061	0,223
27,0 - 28,9	0,204	0,131	0,335
29,0 - 30,9	0,071	0,144	0,215
31,0 - 32,9	0,013	0,098	0,111
33,0 et plus	0,001	0,048	0,049
Ensemble	0,500	0,500	1,000
Moyenne	27,25	29,66	28,44
Écart type	1,77	2,75	2,50
<i>Sexe féminin</i>			
Moins de 30,0	0,032	0,011	0,043
30,0 - 31,9	0,067	0,035	0,102
32,0 - 33,9	0,101	0,082	0,183
34,0 - 35,9	0,096	0,125	0,221
36,0 - 37,9	0,062	0,133	0,195
38,0 - 39,9	0,028	0,104	0,132
40,0 - 41,9	0,010	0,063	0,073
42,0 et plus	0,004	0,047	0,051
Ensemble	0,400	0,600	1,000
Moyenne	34,14	36,86	35,77
Écart type	3,03	3,53	3,57

3. *Maladies cérébro-vasculaires*

Proportion de décès par maladies cérébro-vasculaires (en %)	Fréquences ajustées		
	<i>D</i> <sub>1</sub>	<i>D</i> <sub>2</sub>	Ensemble
	<i>Sexe masculin</i>		
Moins de 7,0	0,058	0,027	0,085
7,0 - 7,9	0,173	0,094	0,267
8,0 - 8,9	0,171	0,149	0,320
9,0 - 9,9	0,071	0,124	0,195
10,0 - 10,9	0,025	0,073	0,098
11,0 et plus	0,002	0,033	0,035
Ensemble	0,500	0,500	1,000
Moyenne	8,15	9,01	8,56
Écart type	1,04	1,40	1,23
	<i>Sexe féminin</i>		
Moins de 8,0	0,031	0,009	0,040
8,0 - 8,9	0,076	0,035	0,111
9,0 - 9,9	0,124	0,077	0,201
10,0 - 10,9	0,124	0,109	0,233
11,0 - 11,9	0,078	0,107	0,185
12,0 - 12,9	0,040	0,078	0,118
13,0 - 13,9	0,018	0,052	0,070
14,0 et plus	0,009	0,033	0,042
Ensemble	0,500	0,500	1,000
Moyenne	10,23	11,31	10,77
Écart type	1,50	1,76	1,76

4. *Maladies du cœur*

Proportion de décès par maladies du cœur (en %)	Fréquences ajustées		
	<i>D</i> <sub>1</sub>	<i>D</i> <sub>2</sub>	Ensemble
	<i>Sexe masculin</i>		
Moins de 14,0	0,021	0,009	0,030
14,0 - 15,9	0,106	0,054	0,160
16,0 - 17,9	0,209	0,121	0,330
18,0 - 19,9	0,162	0,119	0,281
20,0 - 21,9	0,078	0,068	0,146
22,0 - 23,9	0,019	0,022	0,041
24,0 et plus	0,005	0,007	0,012
Ensemble	0,600	0,400	1,000
Moyenne	17,79	18,28	18,05
Écart type	2,29	2,37	2,39
	<i>Sexe féminin</i>		
Moins de 10,0	0,015	0,005	0,020
10,0 - 11,9	0,094	0,058	0,152
12,0 - 13,9	0,160	0,147	0,307
14,0 - 15,9	0,119	0,170	0,289
16,0 - 17,9	0,045	0,112	0,157
18,0 - 19,9	0,013	0,043	0,056
20,0 et plus	0,004	0,015	0,019
Ensemble	0,450	0,550	1,000
Moyenne	13,63	14,86	14,32
Écart type	2,25	2,44	2,50

## 5. Diabète

Proportion de décès par diabète (en %)	Fréquences ajustées		
	$D_1$	$D_2$	Ensemble
<i>Sexe masculin</i>			
Moins de 0,7 . . . . .	0,043	0,001	0,044
0,7 - 0,8 . . . . .	0,132	0,033	0,165
0,9 - 1,0 . . . . .	0,092	0,139	0,231
1,1 - 1,2 . . . . .	0,027	0,213	0,240
1,3 - 1,4 . . . . .	0,005	0,171	0,176
1,5 - 1,6 . . . . .	0,001	0,090	0,091
1,7 et plus . . . . .	ε	0,053	0,053
Ensemble . . . . .	0,300	0,700	1,000
Moyenne . . . . .	0,88	1,29	1,17
Écart type . . . . .	0,17	0,27	0,31
<i>Sexe féminin</i>			
Moins de 1,0 . . . . .	0,017	0,004	0,021
1,0 - 1,4 . . . . .	0,119	0,065	0,184
1,5 - 1,9 . . . . .	0,111	0,177	0,288
2,0 - 2,4 . . . . .	0,041	0,193	0,234
2,5 - 2,9 . . . . .	0,010	0,133	0,143
3,0 - 3,4 . . . . .	0,002	0,071	0,073
3,5 et plus . . . . .	ε	0,057	0,057
Ensemble . . . . .	0,300	0,700	1,000
Moyenne . . . . .	1,61	2,37	2,14
Écart type . . . . .	0,45	0,77	0,77

## 6. Alcoolisme

Proportion de décès par alcoolisme (en %)	Fréquences ajustées		
	$D_1$	$D_2$	Ensemble
<i>Sexe masculin</i>			
Moins de 1,0 . . . . .	0,005	0,001	0,006
1,0 - 1,4 . . . . .	0,045	0,018	0,063
1,5 - 1,9 . . . . .	0,081	0,085	0,166
2,0 - 2,4 . . . . .	0,074	0,149	0,223
2,5 - 2,9 . . . . .	0,047	0,154	0,201
3,0 - 3,4 . . . . .	0,025	0,120	0,145
3,5 - 3,9 . . . . .	0,012	0,078	0,090
4,0 - 4,4 . . . . .	0,006	0,045	0,051
4,5 et plus . . . . .	0,005	0,050	0,055
Ensemble . . . . .	0,300	0,700	1,000
Moyenne . . . . .	2,25	2,96	2,72
Écart type . . . . .	0,82	0,98	0,92
<i>Sexe féminin</i>			
Moins de 0,4 . . . . .	0,013	0,001	0,014
0,4 - 0,7 . . . . .	0,182	0,040	0,222
0,8 - 1,1 . . . . .	0,236	0,097	0,333
1,2 - 1,5 . . . . .	0,144	0,080	0,224
1,6 - 1,9 . . . . .	0,069	0,045	0,114
2,0 - 2,3 . . . . .	0,031	0,021	0,052
2,4 - 2,7 . . . . .	0,014	0,008	0,022
2,8 - 3,1 . . . . .	0,006	0,006	0,012
3,2 et plus . . . . .	0,005	0,002	0,007
Ensemble . . . . .	0,700	0,300	1,000
Moyenne . . . . .	1,17	1,36	1,22
Écart type . . . . .	0,57	0,58	0,57

## 7. Cirrhose du foie

Proportion de décès par cirrhose du foie (en %)	Fréquences ajustées		
	$D_1$	$D_2$	Ensemble
	<i>Sexe masculin</i>		
Moins de 7,0	0,033	0,010	0,043
7,0 - 7,9	0,093	0,052	0,145
8,0 - 8,9	0,131	0,117	0,248
9,0 - 9,9	0,105	0,145	0,250
10,0 - 10,9	0,056	0,115	0,171
11,0 - 11,9	0,022	0,066	0,088
12,0 - 12,9	0,007	0,029	0,036
13,0 et plus	0,003	0,016	0,019
Ensemble	0,450	0,550	1,000
Moyenne	8,86	9,77	9,37
Écart type	1,38	1,53	1,56
	<i>Sexe féminin</i>		
Moins de 3,0	0,016	0,005	0,021
3,0 - 4,9	0,154	0,075	0,229
5,0 - 6,9	0,182	0,150	0,332
7,0 - 8,9	0,095	0,127	0,222
9,0 - 10,9	0,035	0,075	0,110
11,0 - 12,9	0,012	0,037	0,049
13,0 - 14,9	0,004	0,017	0,021
15,0 et plus	0,002	0,014	0,016
Ensemble	0,500	0,500	1,000
Moyenne	6,16	7,80	6,98
Écart type	2,25	3,05	2,82

## 8. Suicide

Proportion de décès par suicide (en %)	Fréquences ajustées		
	$D_1$	$D_2$	Ensemble
	<i>Sexe masculin</i>		
Moins de 2,0	0,012	0,001	0,013
2,0 - 2,9	0,192	0,041	0,233
3,0 - 3,9	0,162	0,177	0,339
4,0 - 4,9	0,031	0,204	0,235
5,0 - 5,9	0,003	0,115	0,118
6,0 - 6,9	ε	0,044	0,044
7,0 et plus	ε	0,018	0,018
Ensemble	0,400	0,600	1,000
Moyenne	3,05	4,50	3,92
Écart type	0,66	1,16	1,24
	<i>Sexe féminin</i>		
Moins de 1,0	0,022	0,014	0,036
1,0 - 1,9	0,208	0,158	0,366
2,0 - 2,9	0,169	0,170	0,339
3,0 - 3,9	0,068	0,090	0,158
4,0 - 4,9	0,023	0,040	0,063
5,0 - 5,9	0,007	0,016	0,023
6,0 et plus	0,003	0,012	0,015
Ensemble	0,500	0,500	1,000
Moyenne	2,29	2,68	2,48
Écart type	1,04	1,31	1,19

## 9. Accidents

Proportion de décès par accidents (en %)	Fréquences ajustées		
	$D_1$	$D_2$	Ensemble
	<i>Sexe masculin</i>		
Moins de 7,0 . . . . .	0,022	0,010	0,032
7,0 - 7,9 . . . . .	0,084	0,034	0,118
8,0 - 8,9 . . . . .	0,154	0,085	0,239
9,0 - 9,9 . . . . .	0,134	0,131	0,265
10,0 - 10,9 . . . . .	0,073	0,103	0,176
11,0 - 11,9 . . . . .	0,024	0,069	0,093
12,0 - 12,9 . . . . .	0,007	0,041	0,048
13,0 et plus . . . . .	0,002	0,027	0,029
Ensemble . . . . .	0,500	0,500	1,000
Moyenne . . . . .	9,01	10,05	9,56
Écart type . . . . .	1,27	1,73	1,59
	<i>Sexe féminin</i>		
Moins de 3,0 . . . . .	0,032	0,001	0,033
3,0 - 3,9 . . . . .	0,121	0,021	0,142
4,0 - 4,9 . . . . .	0,144	0,114	0,258
5,0 - 5,9 . . . . .	0,071	0,200	0,271
6,0 - 6,9 . . . . .	0,023	0,154	0,177
7,0 - 7,9 . . . . .	0,007	0,070	0,077
8,0 et plus . . . . .	0,002	0,040	0,042
Ensemble . . . . .	0,400	0,600	1,000
Moyenne . . . . .	4,39	5,94	5,33
Écart type . . . . .	1,11	1,28	1,43

## CONCLUSION

Cette étude montre l'intérêt que peut présenter l'analyse des statistiques générales à l'aide de modèles.

En étudiant les distributions départementales des proportions de décès par cause, on a trouvé qu'elles étaient la somme de deux distributions log-normales. On en a déduit que, pour une cause de décès donnée, la population totale était composée de deux populations. On a proposé d'expliquer ce résultat par des caractéristiques différentes de la mortalité pour la cause considérée : pour une des populations, la mortalité est d'origine exogène et, pour l'autre, d'origine endogène.

## ANNEXE

## 1. Liaison entre la mortalité relative et le taux de prévalence

On définit les quantités suivantes, pour un sexe et un groupe d'âge donnés et pour le département  $j$  :

- $m_j$  et  $m_{kj}$ , taux de mortalité générale et taux de mortalité de la maladie  $k$ ,
- $n_{kj}$ , nombre de malades par habitant ou taux de prévalence de la maladie  $k$ ,
- $x_{kj}$ , mortalité relative de la maladie  $k$ , c'est-à-dire proportion des décès de la cause  $k$  par rapport à l'ensemble des décès.

On suppose que le taux de mortalité de la maladie  $k$  rapporté au nombre de malades correspondants, qui a pour expression  $m_{kj}/n_{kj}$ , est proportionnel au taux de la mortalité générale :

$$m_{kj}/n_{kj} = \gamma_k m_j$$

où  $\gamma_k$  est un coefficient dépendant de la maladie et indépendant du département.

Comme :  $x_{kj} = m_{kj}/m_j$ , il vient :

$$x_{kj} = \gamma_k n_{kj} \quad (2)$$

Pour l'ensemble de la France, la relation (2) s'écrit :

$$x_k = \gamma_k n_k \quad (3)$$

où  $n_k$  et  $x_k$  sont les taux de prévalence et de mortalité relative de la maladie  $k$  pour la France entière. Les valeurs de  $x_k$  se calculent à partir des statistiques de causes de décès; celles de  $n_k$  ont été évaluées dans une étude précédente [5].

La relation (2) permet donc de calculer les valeurs du coefficient  $\gamma_k$ .

## 2. Méthode d'ajustement

Pour ajuster les distributions, on utilise la méthode de la droite de Henri, en fonction du logarithme de la variable étudiée.

On se donne l'effectif  $N_1$  de la première distribution et on suppose que les effectifs des classes correspondant aux valeurs les plus faibles de  $x$  appartiennent à cette distribution. On construit la droite de Henri correspondante et on en déduit des estimations des fréquences des autres classes.

Par différence avec les fréquences de la distribution observée, on obtient des évaluations des fréquences de la deuxième distribution. Sur ces valeurs, on trace une deuxième droite de Henri à partir de laquelle on calcule les fréquences ajustées de cette distribution. On compare, par la méthode du chi deux, la distribution observée avec la somme des fréquences des deux distributions ajustées.

On essaie plusieurs valeurs de  $N_1$ , on conserve celle pour laquelle la valeur du chi deux est minimale.

## 3. Calcul des caractéristiques de la distribution log-normale

La variable aléatoire  $x$  suit une loi log-normale si la quantité :

$$z = a \text{Log} (x - x_0) + b$$

suit une loi normale réduite.

Dans cette étude,  $x_0 = 0$  et  $z$  s'écrit :

$$z = a \text{Log} x + b = au + b$$

où  $a$  et  $b$  sont des paramètres et  $u = \text{Log} x$  représente le logarithme népérien de  $x$ .

Cette relation linéaire entre  $z$  et  $u$  représente l'équation de la droite de Henri.

On peut écrire :

$$z = \frac{u - \bar{u}}{s_u}$$

d'où les valeurs de la moyenne et de l'écart type de  $u$  :

$$\bar{u} = -\frac{b}{a}, s_u = \frac{1}{a}$$

On démontre que la moyenne et l'écart type de  $x$  ont pour expression :

$$\bar{x} = \exp \left\{ \frac{1}{2a^2} - \frac{b}{a} \right\}$$

$$s_x = \exp \left\{ \frac{1}{2a^2} - \frac{b}{a} + \frac{1}{2} \text{Log} \left( e^{\frac{1}{a^2}} - 1 \right) \right\}$$

On calcule la moyenne et l'écart type de chacune des distributions composantes à l'aide de ces formules. Pour la distribution d'ensemble, on utilise les formules habituelles.

#### 4. Calcul des taux de prévalence et des taux de mortalité rapportés aux malades

On définit, pour la maladie  $k$  et par sexe :

- $x_{k1}, x_{k2}$ , mortalité relative des populations  $P_1$  et  $P_2$ ,
- $n_{k1}, n_{k2}, n_k$ , taux de prévalence des populations  $P_1, P_2$  et de l'ensemble,
- $m'_{k1}, m'_{k2}, m'_k$ , taux de mortalité rapportés au nombre des malades pour les populations  $P_1$  et  $P_2$  et pour l'ensemble,
- $\alpha_{k1}, \alpha_{k2}$ , proportion des effectifs des populations  $P_1$  et  $P_2$  par rapport à l'ensemble ( $\alpha_{k1} + \alpha_{k2} = 1$ ).

Pour les taux de prévalence des populations  $P_1$  et  $P_2$ , on a :

$$n_{k1} = \alpha_1 n_k$$

et :

$$n_{k2} = \alpha_2 n_k$$

Pour les taux de mortalité rapportés au nombre de malades, on a les relations :

$$m'_k = \alpha_{k1} m'_{k1} + \alpha_{k2} m'_{k2}$$

et :

$$\frac{m'_{k1}}{m'_{k2}} = \frac{x_{k1}}{x_{k2}} \frac{\alpha_2}{\alpha_1}$$

On en tire :

$$m'_{k1} = \frac{m'_k}{\alpha_{k1}} \frac{x_{k1}}{x_{k1} + x_{k2}}$$

et :

$$m'_{k2} = \frac{m'_k}{\alpha_{k2}} \frac{x_{k2}}{x_{k1} + x_{k2}}$$

Les valeurs de  $x_{k1}, x_{k2}, \alpha_{k1}, \alpha_{k2}$  ont été calculées dans la présente étude. Les valeurs de  $m'_k$  sont tirées de l'étude [5].

#### RÉFÉRENCES

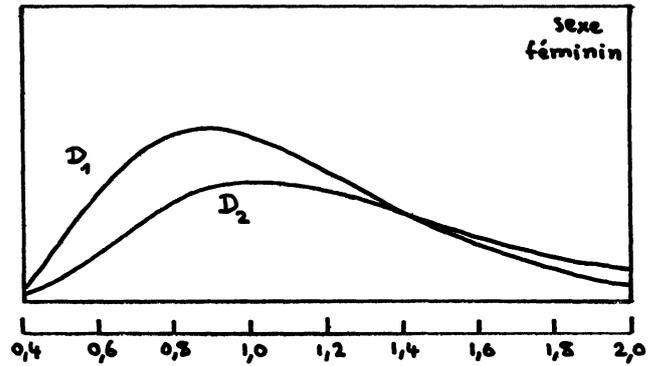
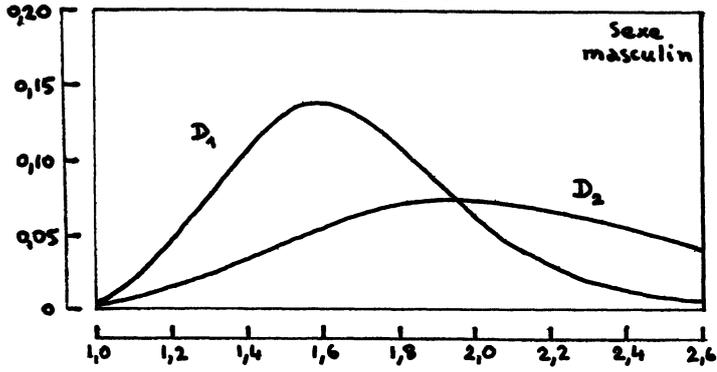
- [1] DAMIANI Paul, MASSÉ Hélène, AUBENQUE Maurice. — Étude des distributions de consommation de tabac et d'alcool. *Journal de la Société de statistique de Paris*, tome 124, n° 2, 1983, 119-128.
- [2] Statistique des causes médicales de décès. Volumes annuels. I.N.S.E.R.M.
- [3] BREIL Jacques. — Statistique du mouvement de la population, 2<sup>e</sup> partie : Les causes de décès. Année 1943, p. XI-XV, 1947.
- [4] LEDERMANN Sully. — La répartition des décès de cause indéterminée. *Revue de l'Institut international de statistique*, I-III, 1956, 47-57.
- [5] DAMIANI Paul, MASSÉ Hélène, AUBENQUE Maurice. — Évaluation de la morbidité à partir de la mortalité. *Journal de la Société de statistique de Paris*, tome 123, n° 1, 1982, 58-74.

GRAPHIQUES

Distribution des proportions départementales de mortalité par cause,  
suivant le sexe, pour le groupe d'âge 45-64 ans

1. Tuberculose

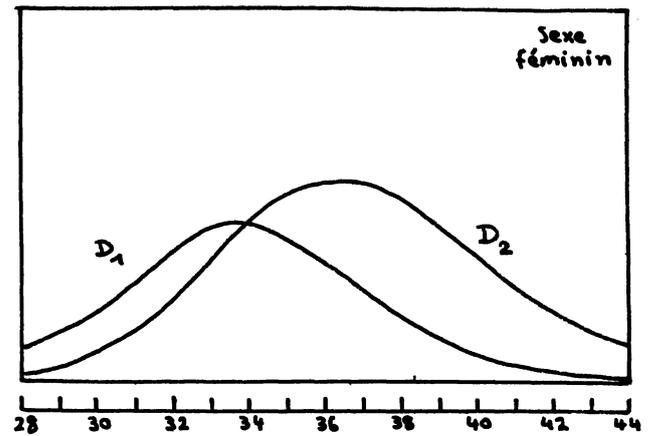
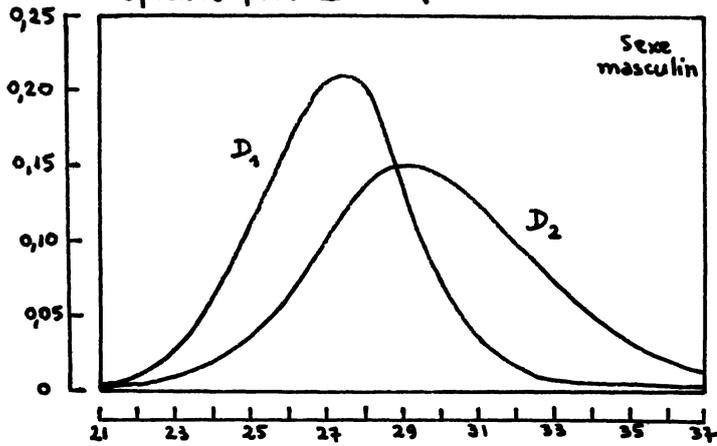
Fréquence pour  $\Delta x = 0,2$  p.100



x, proportion pour cent de mortalité par tuberculose

2. Cancers

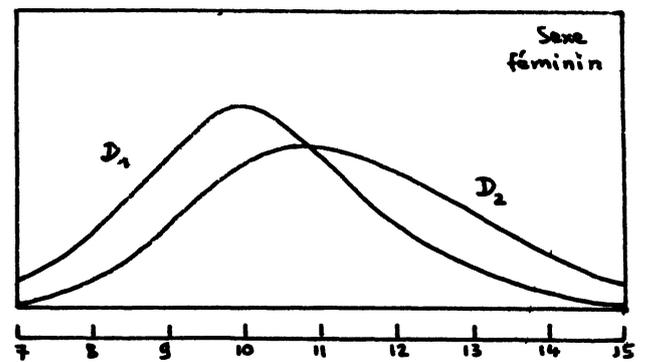
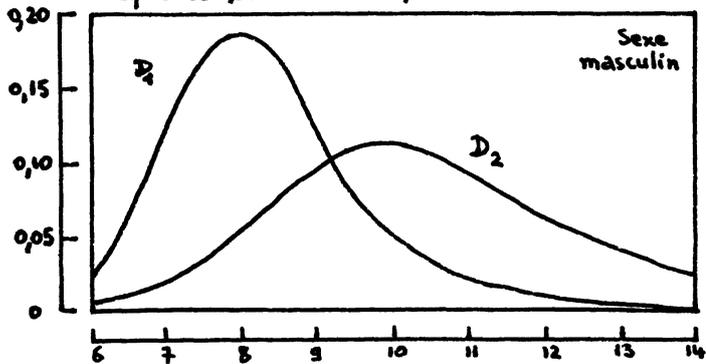
Fréquence pour  $\Delta x = 2$  p.100



x, proportion pour cent de mortalité par cancers

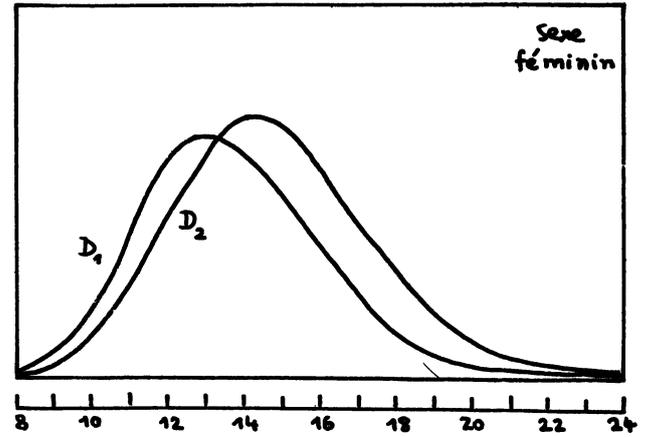
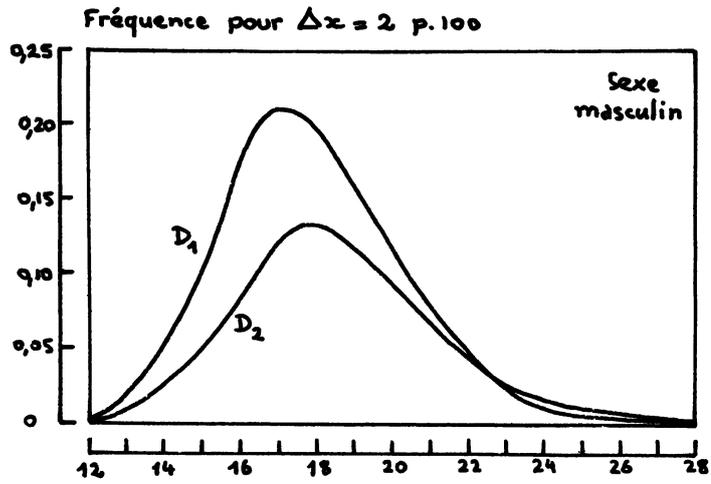
3. Maladies cérébro-vasculaires

Fréquence pour  $\Delta x = 1$  p.100



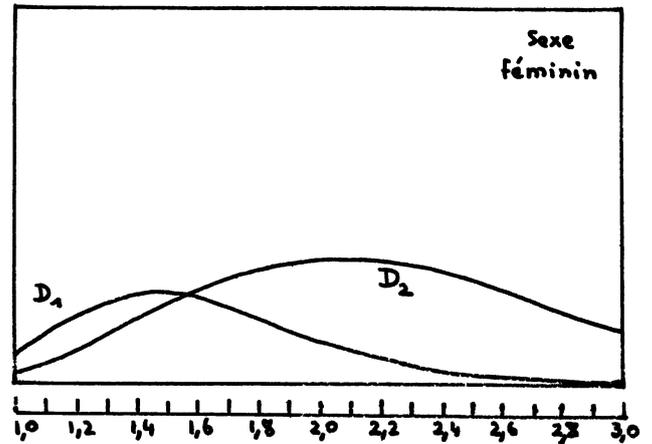
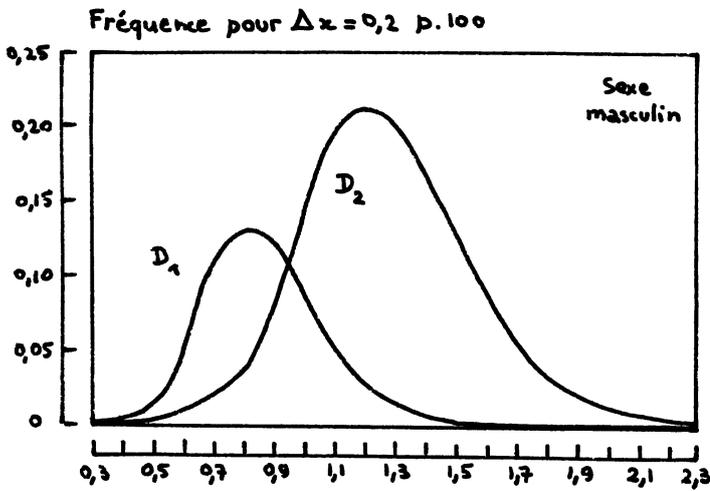
x, proportion pour cent de mortalité par maladies cérébro-vasculaires

#### 4. Maladies du cœur



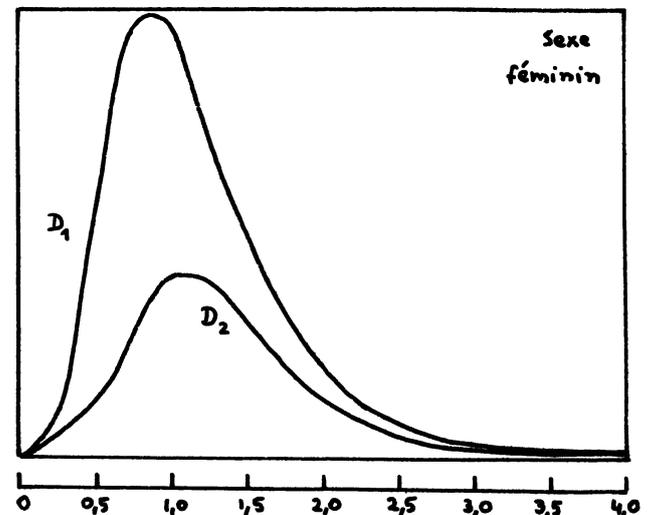
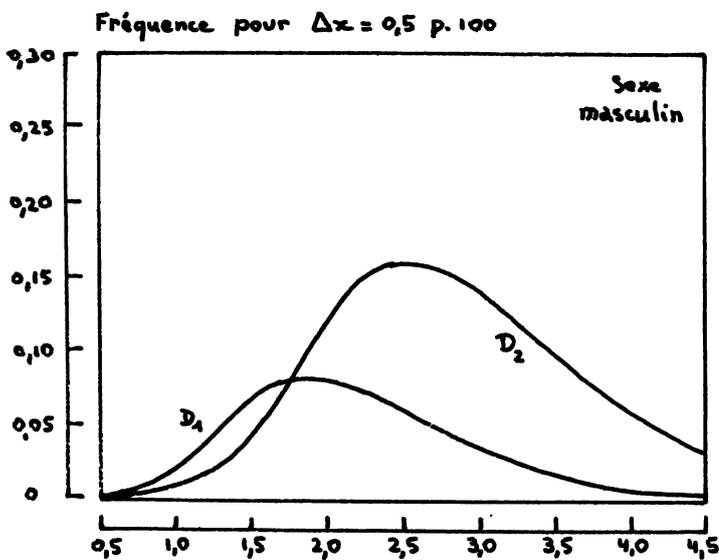
x, proportion pour cent de mortalité par maladies du cœur

#### 5. Diabète



x, proportion pour cent de mortalité par diabète

#### 6. Alcoolisme



x, proportion pour cent de mortalité par alcoolisme

