Sur l'ajustement des tables de mortalité pour les âges auxquels la formule de Makeham n'est pas applicable

Par *Ch. Willigens*, Dr ès sciences mathématiques à Berne

Le présent travail peut être considéré comme une suite à la note que j'ai publiée dans le temps sur l'ajustement des tables de mortalité de la population suisse pour les années 1920 et 1921 (Journal de statistique et revue économique suisse, 1926, p. 339). Dans cette première note, j'ai étudié l'ajustement par la formule de Makeham en opérant sur les taux de mortalité annuels. Cette méthode présente l'avantage de permettre l'usage d'une machine à calculer pour les nombres de huit chiffres seulement. Je me suis contenté pour les premiers âges d'un ajustement mécanique. Depuis j'ai été conduit à une formule de représentation qui est valable à partir de l'âge de 5 ans pour le sexe masculin, à partir de 3 ans pour le sexe féminin.

Supposons qu'une table ait été ajustée par la formule de Makeham. Si l'on suppose connu l'ordre de survie définitif et si l'on désigne par l_x le nombre réel de survivants d'âge x, par λ_x le nombre fourni par la formule de Makeham, on a

$$l_x \, a \, \lambda_x$$

à partir du premier âge auquel la formule de Makeham est appliquable, jusque vers les premières années d'âge. Le rapport $\frac{l_x}{\lambda_x}$ part donc de la valeur 1, et passe par un minimum pour revenir ensuite à la valeur 1 on est donc conduit à tenter de représenter ce rapport par la formule

$$\frac{l_x}{\lambda_x} = 1 - \varrho \, (x - \sigma)^\alpha \, e^{-\nu x}.$$

Le second terme du second membre est nul pour $x = \sigma$, passe par un maximum et tend asymptotiquement vers zéro. Au début de la validité de la formule de Makeham il sera pratiquement très voisin de la valeur zéro, étant donné le degré d'approximation cherché, et n'aura plus qu'une très faible influence sur les premiers termes de cet ajustement. Vu l'ordre de grandeur des nombres qui interviennent, le carré de ce nombre est d'ordre inférieur à celui de la cinquième décimale.

On peut donc poser

$$\frac{l_x}{\lambda_x} = 1 - \varrho \, (x - \sigma)^\alpha \, e^{-\nu x} = e^{-\varrho \, (x - \sigma)^\alpha \, e^{-\nu x}}.$$

Posons $-\varrho = \text{Log.}\, f \quad -\nu = \text{Log.}\, h$ la formule prend la forme

$$l_x = K\, s^x g^{c^x}\, f^{h^x \cdot (x-\sigma)^a}$$

en remplaçant γ_x par l'expression de la formule de Nakeham. Dans le cas particulier de $a = o$ on retrouve la formule connue de Lazarus.

Avant de passer à la méthode d'ajustement proprement dite, il y a lieu de rappeler comment Pearson détermine les coefficients pour une suite de nombres représentable par la formule

$$y = \varrho\, (x-\sigma)^a\, e^{-\nu x}.$$

Cette expression satisfait à une équation différentielle de la forme

$$\frac{1}{y}\,\frac{dy}{dx} = \frac{a - a}{b_0 + b_1\, x}$$

où

$$a = \frac{a}{\nu} + \sigma$$

$$b_0 = \frac{\sigma}{\nu}$$

$$b_1 = \frac{1}{\nu}.$$

Multiplions les deux membres de l'équatron différentielle par x^ν après avoir chassé les dénominateurs. En intégrant le premier membre par parties, il vient:

$$\left[x^\nu\, y\, (b_0 + b_1\, x)\right]_{x_0}^{\infty} - \int_{x_0}^{\infty} y\, \frac{d}{dx}\left[x^\nu (b_0 + b_1 x)\right] dx = \int_{x_0}^{\infty} x^\nu\, (x-a)\; y\; dx.$$

On peut supposer que le crochet au 1er membre est nul pour $x = x_0$, il est également nul pour $x = \infty$. D'autre part, on peut évaluer numériquement les intégrales

$$\mu_\nu = \int_a^{\infty} x^\nu\; y\; dx.$$

En donnant à v les valeurs 0, 1 et 2, on obtient les trois relations suivantes qui permettent de déterminer b_0, b_1 et a et par suite a, σ et ν:

$$b_1\, \mu_0 = a\, \mu_0 - \mu_1$$

$$b_0\, \mu_0 + 2 b_1\, \mu_1 = a\, \mu_1 - \mu_2$$

$$2 b_0\, \mu_1 + 3 b_1\, \mu_2 = a\, \mu_2 - \mu_3.$$

Il pourra être commode de déterminer directement a, qui est l'abscisse du maximum, par interpolation à l'aide d'une parabole dont on détermine le sommet et de déterminer b_0 et b_1 à l'aide des deux premières équations.

Pratiquement j'ai trouvé qu'il était plus avantageux de déterminer seulement la valeur de σ à l'aide de ces équations et, une fois σ connu de, déterminer les constantes a, ν et ϱ par une méthode analogue à celle de King et Hardy pour la formule de Makeham. On peut, bien entendu, se servir également de la méthode des moindres carrés.

Ceci posé, désignons par π_x la probabilité de survie qui résulte de la formule de Makeham. Nous obtenons pour le rapport $\frac{p_x}{\pi_x}$ en retenant seulement les deux premiers termes du développement en série:

$$\frac{p_x}{\pi_x} = 1 - \varrho\left[(x+1-\sigma)^{\alpha} e^{-\nu(x+1)} - (x-\sigma)^{\alpha} e^{-\nu x}\right]$$

En faisant l'approximation admissible dans le cas présent

$$\varphi(x+1) - \varphi(x) = 1 \times \varphi^{1}(x)$$

nous obtenons

$$\frac{p_x}{\pi_x} = 1 + \varrho\nu(x-\sigma)^{\alpha-1} e^{-\nu x}\left(x - \frac{\alpha}{\nu} - \sigma\right)$$

donc

$$\frac{p_x}{\pi_x} - 1 = \varrho\nu(x-\sigma)^{\alpha-1} e^{-\nu x}(x-i)$$

i correspond à l'abscisse du point d'intersection des courbes représentatives de p_x et de π_x. Le nombre i peut être déterminé directement par interpolation. Si la courbe des p_x est trop irrégulière dans le voisinage du point d'intersection, il sera indiqué de procéder à un ajustement mécanique dans le voisinage de ce point, avant de calculer i.

On a donc à calculer la suite de nombres

$$\left(\frac{p_x}{\pi_x} - 1\right) \times \frac{1}{x-i} = \varrho\nu(x-\sigma)^{\alpha-1} e^{-\nu x}$$

et l'on procède à l'ajustement de ces valeurs par la formule indiquée dans le second membre. Les valeurs ajustées étant connues, on multiplie chacune par la valeur correspondante de $x-i$, ce qui donne les valeurs $\frac{p_x}{\pi_x} - 1$ d'où l'on déduit p_x ajusté.

Je me contenterai de donner les valeurs ajustées et non ajustées des taux de mortalité q_x et les valeurs ajustées par la formule de Makeham pour les âges où elles sont valables. Ces derniers nombres sont en italiques dans le tableau qui va jusqu'à l'âge de 30 ans. Les tables complètes, ainsi que les commutations, ont été publiées dans le rapport du bureau fédéral des assurances pour l'année 1925, mais avec un ajustement mécanique pour le début de la table.

Taux de mortalité des tables SM et SF 1920/21

x	sexe masculin		sexe féminin		x	sexe masculin		sexe féminin	
	non ajustés et Makeham	ajustés	non ajustés et Makeham	ajustés		non ajustés et Makeham	ajustés	non ajustés et Makeham	ajustés
0	0,09051	0,09051	0,07016	0,07016	15	256	285	268	295
1	1593	1593	1455	1455	16	294	309	336	324
2	656	656	711	711	17	356	330	359	351
3	543	543	495	502	18	*376*	349	375	374
4	399	399	411	428	19	*382*	364	377	395
5	315	346	369	388	20	*389*	377	418	411
6	322	323	315	332	21	*396*	388	401	426
7	300	285	245	273	22	*405*	400	444	438
8	227	247	257	224	23	*414*	411	*465*	448
9	218	217	229	193	24	*423*	422	*469*	457
10	216	201	198	182	25	*434*	434	*474*	465
11	220	200	202	188	26	*445*	445	*479*	473
12	210	212	188	207	27	*458*	458	*484*	480
13	214	233	210	233	28	*472*	472	*490*	487
14	218	258	247	263	29	*487*	487	*497*	495
					30	*504*	504	*505*	504

Les valeurs des constantes qui interviennent dans la formule

$$l_x = K\, s^x\, g^{c^x}\, e^{-\varrho\,(x-\sigma)^a}\, e^{-\nu x}$$ sont:

Sexe masculin.

$K = 91.030,4$	$\sigma = -6,105$
$s = 0,99.690$	$a = 16,50.921$
$g = 0,998572$	$\nu = 1,11.582$

Sexe féminin.

$K = 94.533,2$	$\sigma = +0,54.465$
$s = 0,99.570$	$a = 4,75.630$
$g = 0,999742$	$\nu = 0,54.342$
$c = 1,114$	$\log \varrho\nu = +0,30.885 - 5$

Cette formule n'est, bien entendu, qu'une expression donnant une bonne interpolation de la table, je n'ai nullement la prétention d'avoir découvert une loi naturelle.