Université Paris 1 Panthéon Sorbonne,

Institut de démographie

Chapitre 6 :

Analyse des interférences : tables d'extinction multiple

1. Evénements perturbateurs et événements concurrents
2. Estimation d'une probabilité « nette » et d'une probabilité « brute ».
3. Présentation des événements concurrents (graphique et table combinée)
4. Construction et application des table « brutes » pour analyser des événements concurrents
5. Construction des tables « nettes » (supposant l'absence des événements concurrents)
6. Tables associées à une seule cause de sortie (de décès)

Lecture (avec précaution) : H.Léridon et L.Toulemon *Démographie. Approche statistique et dynamique de la population*. Paris, 1997, Economica, chapitres 7 et 8 (p.92-137)

Cours d'analyse démographique par Alexandre Avdeev, niveau : **Master 1e année et Diplôme générale de démographie**

Les événements concurrents et perturbateurs

Les événements concurrents: l'arrivée E_1 empêche E2 de se produire, et réciproquement (symétrie)

- Décès par cause
- Décès et émigration (l'émigration empêche le décès sur place)
- Mariage et décès des célibataire
- Divorce et veuvage

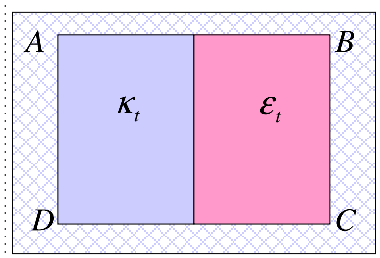
Les événements perturbateurs: l'arrivée de E2 empêche E1, mais la réciproque n'est pas vrai

- Mariage et émigration (l'émigration empêche le mariage sur place)
- Naissance dans le mariage et divorce où veuvage

Les probabilités « brutes »

Soit

$$\kappa_t = \frac{D_t}{P_t} \rightarrow$$ probabilité de mourir $$\varepsilon_t = \frac{E_t}{P_t} \rightarrow$$ probabilité de partir



Présentation de des probabilités brutes sur un diagramme de Carroll*) :

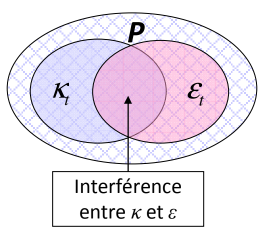
- soit ***s*** la surface du rectangle ABCD représentant la probabilité de sortir de l'observation en raison du décès ou du départ ;
- la surface du rectangle ABCD est la somme des surfaces d'un rectangle $\boldsymbol{\kappa}$ et d'un rectangle $\boldsymbol{\varepsilon}$

$$s = \frac{D_t}{P_t} + \frac{E_t}{P_t} = \frac{D_t + E_t}{P_t} \rightarrow \textit{probabilité de sortir de la table}$$

Les probabilités brutes sont additives.

*) Lewis Carroll = Charles Lutwidge Dodgson (1832-1898)

Les probabilités « nettes »



Présentation de des probabilités nettes sur un diagramme de d'Euler - Venn :

- soit ***P*** la surface de l'ellipse représentant la probabilité de tout événement possible (probabilité complète = 1) ;
- la probabilité de sortir de l'observation n'est pas égale à la somme des surfaces d'un cercle $\boldsymbol{\kappa}$ (probabilité « nette » de mourir) et d'un cercle $\boldsymbol{\varepsilon}$ *(probabilité « nette » de partir)*

formule de Berkson

$$q_t = \frac{D_t}{P_t - 0{,}5 \cdot E_i} \rightarrow$$ probabilité de mourir en absence de l'émigration
avec une hypothèse de « fifty-fifty » (50% part au début et 50% à la fin de la période)

$$e_t = \frac{E_t}{P_t - 0{,}5 \cdot D_i} \rightarrow$$ probabilité de partir en absence de la mortalité

<u>Les probabilités nettes ne sont pas additives,</u> mais leur compléments sont conditionnels :

$$1 - s_t = (1 - q_t) \cdot (1 - e_t)$$

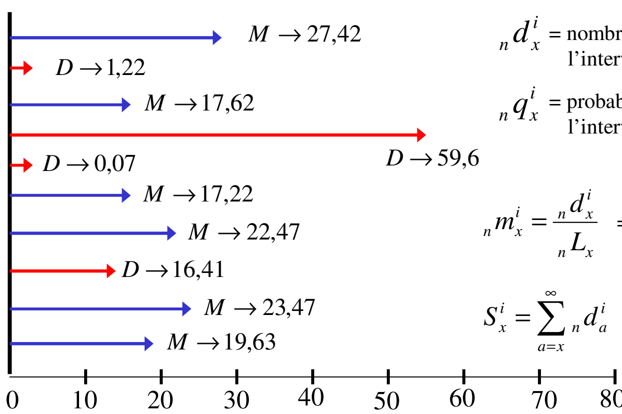
Lecture: la probabilité de rester dans la table est égale la probabilité de ne pas mourir multipliée par la probabilité de ne pas partir

donc $s_t = 1 - (1 - q_t) \cdot (1 - e_t)$ $\quad s_t = q_t + e_t - q_t \cdot e_t$

Tables des risques combinés et tables de risque net

- ***Tables des risques combinés*** représentent un modèle d'extinction d'une population à cause de plusieurs facteurs de manière que la contribution de chacun entre eux ne peut être mesure en absence des contributions des autres facteurs.
- ***Tables de risque net*** représentent un modèle d'évolution d'une population soumise à l'influence des plusieurs facteurs d'extinction de manière que la contribution de chaque facteur est mesurée comme si les autres sont totalement absents.

Présentation des événements concurrents (données individuelles)



${}_n d_x^i$ = nombre de sortants par cause i dans l'intervalle x, x+n

${}_n q_x^i$ = probabilité de sortir par cause i dans l'intervalle x, x+n

$${}_n q_x^i = \frac{{}_n d_x^i}{S_x}$$

${}_n m_x^i = \frac{{}_n d_x^i}{{}_n L_x}$ = taux de sortir par cause i dans l'intervalle x, x+n

$S_x^i = \sum_{a=x}^{\infty} {}_n d_a^i$ = nombre de survivants à l'âge exact x, qui sortiront éventuellement par cause i

$$\sum_i {}_n d_x^i = {}_n d_x \qquad \sum_i {}_n m_x^i = \sum_i \frac{{}_n d_x^i}{{}_n L_x} = \frac{{}_n d_x}{{}_n L_x} = {}_n m_x \qquad \sum_i {}_n q_x^i = \sum_i \frac{{}_n d_x^i}{{}_n S_x} = \frac{{}_n d_x}{{}_n S_x} = {}_n q_x$$

$$S_x^i = \sum_{a=x}^{\infty} {}_n d_a^i \qquad \sum_i S_x^i = \sum_i \sum_{a=x}^{\infty} {}_n d_a^i = \sum_{a=x}^{\infty} {}_n d_a = S_x$$

Une table combinée de nuptialité et de mortalité

Age	Nombre de céliba-taires survi-vants à l'âge ***x***	Nombre de décès des céliba-taires dans l'intervalle d'âge [x,x+n)	Nombre de mariage dans l'inter-valle d'âge [x,x+n)	Nombre de sorties de la table dans l'inter-valle d'âge [x,x+n)	Probabilité de mourir céliba-taires dans l'inter-valle d'âge [x,x+n)	Probabilité de se marier dans l'inter-valle d'âge [x,x+n)	Probabilité de quitter l'état de célibat dans l'intervalle d'âge [x,x+n)	Nombre de célibataire à l'âge ***x*** qui éventu-ellement termine-ront le célibat par décès	Nombre de célibataire à l'âge ***x*** qui éventu-ellement termineront le célibat par mariage	Années vécues dans l'état de célibat dans l'inter-valle d'âge [x,x+n)
x	S_x	${}_nd_x^D$	${}_nd_x^M$	${}_nd_x = {}_nd_x^D + {}_nd_x^M$	${}_nq_x^D$	${}_nq_x^M$	${}_nq_x = {}_nq_x^D + {}_nq_x^M$	S_x^D	S_x^M	${}_nL_x$
0	10	1	0	1	1/10	0	1/10	4	6	9,07
1	9	1	0	1	1/9	0	1/9	3	6	32,2
5	8	0	0	0	0	0	0	2	6	40,0
10	8	1	3	4	1/8	3/8	4/8	2	6	70,88
20	4	0	3	3	0	3/4	3/4	1	3	23,36
30	1	0	0	0	0	0	0	1	0	10,00
40	1	0	0	0	0	0	0	1	0	10,00
50	1	1	0	1	1/1	0	1/1	1	0	9,60
60	0							0	0	

On ne s'intéresse pas aux taux, mais plutôt aux nombre d'années vécues dans l'état de célibat

Table « brute » pour une période:

Hypothèse de base → les événements sont indépendants et complémentaires , donc les probabilité sont additives

On calcule les taux et la probabilité de sortie comme dans une table de mortalité pour une génération :

$${}_nm_x^i = \frac{{}_nd_x^i}{{}_nL_x} \rightarrow \text{ taux et } {}_nq_x^i = \frac{{}_nd_x^i}{S_x} \rightarrow \text{ quotient (probabilité)}$$

comme $$S_x = \frac{1}{n}\cdot\left[{}_nL_x + (n - {}_na_x)\cdot {}_nd_x\right] \quad \textbf{(F1)}$$

la conversion des taux en probabilité nous donne: $${}_nq_x^i = \frac{n\cdot {}_nm_x^i}{1+(n-{}_na_x)\cdot {}_nm_x}$$

Puisque sur l'intervalle d'âge [x, x+n) les événements sont indépendants:

$${}_nm_x = {}_nm_x^i + {}_nm_x^{-i} \Longrightarrow {}_nq_x^i = \frac{n\cdot {}_nm_x^i}{1+(n-{}_na_x)\cdot({}_nm_x^i + {}_nm_x^{-i})} \quad \textbf{(F2)}$$

Relation entre les taux et les probabilités

Une propriété important de **(F2)** $_nq_x^i = \dfrac{n \cdot {}_nm_x^i}{1+(n-{}_na_x)\cdot({}_nm_x^i + {}_nm_x^{-i})}$

si $_nm_x^i = \dfrac{_nd_x^i}{_nL_x} \rightarrow \text{constant} \Rightarrow$ plus grand est $_nm_x^{-i}$ ➡ plus petit est $_nq_x^i$

On considère donc $_nq_x^i$ comme ***la probabilité dépendante***

Par contre, la diminution de $_nm_x^{-i}$ n'a aucune influence sur $_nm_x^i$
l'augmentation du nombre de sortie par cause i est compensé par l'augmentation de nombre d'années d'exposition au risque dans l'intervalle [x, x+1)

Exemple: le taux de mortalité par cancer ne croît pas forcement, si le taux de mortalité les par autres cause se diminue

Calculs d'une table combinée d'extinction multiple

(F2) $_nq_x^i = \dfrac{n \cdot {}_nm_x^i}{1+(n-{}_na_x)\cdot {}_nm_x}$ en divisant $_nq_x^i$ par $_nq_x$ on obtient

$$\frac{_nq_x^i}{_nq_x} = \frac{_nd_x^i}{_nd_x} = \frac{_nm_x^i}{_nm_x} \Rightarrow {}_nq_x^i = {}_nq_x \cdot \frac{_nd_x^i}{_nd_x} = {}_nq_x \cdot \frac{_nm_x^i}{_nm_x}$$

Si nous acceptons une hypothèse que:

$_nm_x^i = {}_nM_x^i \quad et \quad {}_nm_x = {}_nM_x$

où ***M*** – taux observé

$$_nq_x^i = {}_nq_x \cdot \frac{_nM_x^i}{_nM_x} = {}_nq_x \cdot \frac{_nD_x^i}{_nD_x}$$

A partir d'une table de mortalité et de la distribution des décès par cause et par âge on peut calculer ***facilement*** la table d'extinction multiple

1) $_nq_x^i = {}_nq_x \cdot \dfrac{_nD_x^i}{_nD_x}$

2) $_nd_x^i = {}_nq_x^i \cdot S_x$

3) $S_x^i = \sum_{a=x}^{\infty} {}_nd_a^i$

$S_x^i \rightarrow$ Le nombre de survivant à l'âge x qui mourront par cause i

Exemple: Mortalité de tumeur (homme, France, 1999 données modifiées)

Age exact x	Décès nDx	D. de cancer nD^ix	S_x	$_nq_x$	$_nq_x^i$	$_nd_x^i$	S_x^i
0	1 841	22	100000	0.00490	0.00006	6	**32 488**
1	403	73	99510	0.00108	0.00019	19	32 482
5	266	48	99403	0.00071	0.00013	13	32 463
10	349	31	99332	0.00091	0.00008	8	32 450
15	1 446	130	99242	0.00359	0.00032	32	32 442
20	2 126	159	98886	0.00563	0.00042	42	32 410
25	2 419	181	98329	0.00574	0.00043	42	32 368
30	2 806	281	97765	0.00661	0.00066	65	32 326
35	3 981	796	97119	0.00924	0.00185	179	32 261
40	6 506	1 952	96222	0.01539	0.00462	444	32 082
45	9 944	4 176	94741	0.02356	0.00989	937	31 638
50	12 940	6 276	92509	0.03346	0.01623	1501	30 701
55	12 934	6 273	89414	0.04684	0.02272	2031	29 200
60	18 688	9 344	85226	0.06963	0.03481	2967	27 169
65	27 814	12 516	79292	0.10511	0.04730	3750	24 202
70	35 939	15 094	70958	0.15812	0.06641	4712	20 452
75	43 859	15 351	59738	0.23613	0.08265	4937	15 740
80	27 999	8 400	45632	0.36746	0.11024	5030	10 803
85	62 504	12 501	28864	1.00000	0.20000	5773	5 773
Total	274 764	93 605				32 488	

1) $_nq_x^i = {_nq_x} \cdot \frac{_nD_x^i}{_nD_x}$

2) $_nd_x^i = {_nq_x^i} \cdot S_x$

3) $S_x^i = \sum_{a=x}^{\infty} {_nd_a^i}$

Proportion de nouveau-nés qui mourront éventuellement de tumeur:
32 488/100 000=32,5%

Proportion d'homme survécus jusqu'à l'âge de 75 ans qui mourront éventuellement de tumeur:
15 740/59 738=26,3%

Table des risques combinées : un exemple d'étude de la mortalité intra-utérine

- Facteur de risque : la durée de grossesse
- Événement d'intérêt : un décès intra-utérine (fausse couche)
- Événement concurrent : une naissance vivante
- NB : dans le cas de cette études on ne peut pas imaginer une situation dans laquelle risque d'un événement d'intérêt pourrait se réaliser en absence du risque d'un événement concurrent (une population où il n'y a que des fausses couches) et inversement (une population où il n'y a pas de risque de fausse couche)
- Événements perturbateurs :
  - les entrées en observation (troncature à gauche : une femmes avec une durée de grossesse ne peut entrer en observation que si elle n'a pas eu une fausse couche) ;
  - les sorties de l'observation (troncature à droit : si une femme est sortie de l'observation, on ne pourra pas connaître comment sa grossesse se termine-t-elle)

Table des risques combinées : étude de la mortalité intra-utérine (+ un effet de troncature ou des données censurées)

Evénements concurrents : naissances et fausses couches (décès intra-utérine)

Evénement perturbateur : entrée en observation (troncature à gauche) et sortie de l’observation (troncature à droite)

Semaines	Entrées	Fausses couches (décès)	Sorties	Naissances vivantes	Grossesses en cours au début de l'intervalle	Effective au risque	Quotients		Table de mortalité		
							Naissances vivantes	Décès	Survivants	Décès	Nais. Vivantes
t	Et	Dt	St	Nt	(Et-Dt-St-Nt)= Gt	Gt+(Et-St)/2= P_R	v(t)	q(t)	S(t)	d(t)	n(t)
4	592	32	0	0	0	0+(592-0)/2=296	0	0.1081	1000	108.11	0
8	941	72	1	0	592-32-0-0=560	1030	0	0.0699	891.9	62.346	0
12	585	77	2	0	1428	1719.5	0	0.0448	829.5	37.147	0
16	337	28	2	0	1934	2101.5	0	0.0133	792.4	10.558	0
20	248	20	9	1	2241	2360.5	0.0004	0.0085	781.8	6.6244	0.3312
24	175	8	6	4	2459	2543.5	0.0016	0.0031	774.9	2.4372	1.2186
28	98	8	4	25	2616	2663	0.0094	0.003	771.2	2.3169	7.2402
32	67	8	6	72	2677	2707.5	0.0266	0.003	761.7	2.2506	20.255
36	40	9	3	1074	2658	2676.5	0.4013	0.0034	739.2	2.4855	296.61
40	0	11	0	1601	1612	1612	0.9932	0.0068	440.1	3.003	437.07
Total	3083	273	33	2777	Risques basés sur les décès de table →					237.28	762.72
Nb de grossesses observées = 3050 (2777+273)			Risque d’une fausse couche à partir ses observations « directes » (273 : 3050 = 0,0895)→							8.95%	91.05%

Exemple emprunté de Leridon, Toulemon, 1997, p.108-109

Analyse originale F.E.French, J.E.Bierman (1962) « Probablility of fœtal mortality » Public Health Reports, 77 (10)

Table de risque net : un exemple d’étude sur la mortalité de pilotes de Formule 1

- Facteur de risque : l’âge[1)]
- Événements d’intérêt : un décès en compétitions et un décès hors compétition
- Événements concurrents : décès en compétitions et décès hors compétition
- NB : dans le cas de cette études on peut très bien imaginer soit une absence totale de la mortalité associée au risque d’un décès en compétition (les pilotes sont comme tous les autres); soit une absence de la mortalité hors compétition pour estimer le risque net de décès en compétition.
- Événements perturbateurs :
  - les entrées en observation (troncature à gauche : on ne peut devenir un pilote qu’en étant vivant) ;
  - les sorties de l’observation (troncature à droit : on peut décéder hors compétition après avoir quitter une carrier de pilote de F1)

1) Il est possible (et même plus juste d’associer le risque de décès en compétition avec la durée de l’exposition à ce risque (la durée de carrière), toutefois il sera difficile d’y associer le risque de décès hors compétition. Il est possible d’en trouver une solution plausible avec le recours à un modèle de risque proportionnels de D. Cox (proportional hazards model) et en utilisant l’âge au début de carrière comme une variable explicative fixe. La procédure standard de telle analyse est disponible dans tous les systèmes d’analyse statistique (SAS, SPSS, STATA, Statistica etc.)

Table « nette » : mortalité des pilotes

Exemple emprunté de Leridon, Toulemon, 1997, p.103-107
Analyse originale J.-L.Rallu, (1991) « Sélection, âge et performance des coureurs de formule 1 » *Population*, 46,6, p.1711-1733

Données (entrées échelonnées, sorties multiples):

Age	Entrées	Sorties	Décès	Décès en compétition	Décès hors compétition
x	e(x;x+5)	s(x;x+5)	d(x;x+5)	$d^c(x;x+5)$	$d^a(x;x+5)$
20	10	0	0	0	0
25	100	20	4	4	0
30	206	18	9	8	1
35	10	40	7	6	1
40	0	136	4	3	1
45	0	88	0	0	0

Hypothèse 1 :
la densité de risque est uniforme à l'intérieur de chaque intervalle d'âge

Population à la fin de l'intervalle (P_{x+5}):

$$P_{x+5} = P_x + {}_5e_x - {}_5s_x - {}_5d_x$$

Hypothèse 2 :
la moitié d'événements concurrents se font en début et l'autre moitié à la fin de l'intervalle. Alors la population initiale au risque (P_x^R) :

$$P_x^R = P_x + 0{,}5 \cdot {}_5e_x - 0{,}5 \cdot {}_5s_x$$

$${}_5q_x = \frac{{}_5D_x}{P_x + 0{,}5 \cdot {}_5e_x - 0{,}5 \cdot {}_5s_x}$$

$${}_5d_x = S_x \cdot {}_5q_x \qquad S_{x+5} = S_x - {}_5d_x$$

Table de mortalité générale (toutes causes confondues)

x	P_x	P_{x+5}	Population au risque	${}_5D_x$	${}_5q_x$ x1000	${}_5d_x$	S_x
20	0	10	(0+10/2-0/2)= 5	0	0.0	0	1000
25	10	86	(10+100/2-20/2)=50	4	80.0	80	1000
30	86	265	180	9	50.0	46	920
35	265	228	250	7	28.0	24.5	874.0
40	228	88	160	4	25.0	21.2	849.5
45	88	0	44	0	0.0	0.0	828.3

La probabilité de décéder (toutes causes confondues) durant la carrier d'un pilote de formule 1 est :

$${}_{25}q_{20} = 1 - (828{,}3/1000) = 0{,}1717$$

suite

Table de mortalité nette en compétition

x	P_x	P_{x+5}	Population au risque	${}_5D_x$	${}_5q_x$	${}_5d^C{}_x$	$S^C{}_x$
20	0	10	5	0	0.0	0.0	1000
25	10	86	50	4	80.0	80.0	1000
30	86	265	179.5	8	44.6	41.0	920.0
35	265	228	249.5	6	24.0	21.1	879.0
40	228	88	159.5	3	18.8	16.1	857.9
45	88	0	44	0	0.0	0.0	841.7

$${}_{25}q_{20}^C = 1 - (841{,}7/1000) = 0{,}1583$$

$$P_x^R = P_x + 0{,}5 \cdot ({}_5e_x - {}_5s_x - \boxed{{}_5d_x^A})$$

la population au risque est ajustée aux décès hors compétitions

$${}_5q_x^C = \frac{{}_5d_x^C}{P_x + 0{,}5 \cdot ({}_5e_x - {}_5s_x - {}_5d_x^A)}$$

$${}_5d_x^C = S_x^C \cdot {}_5q_x^C$$

$$S_{x+5}^C = S_x^C - {}_5d_x^C$$

Table de mortalité nette hors compétition

x	P_x	P_{x+5}	P^R	${}_5D_x$	${}_5q_x$	${}_5d^A{}_x$	$S^A{}_x$
20	0	10	5	0	0.0	0.0	1000.0
25	10	86	48	0	0.0	0.0	1000.0
30	86	266	176	1	5.7	5.7	1000.0
35	265	229	247	1	4.0	4.0	994.3
40	228	89	158.5	1	6.3	6.2	990.3
45	88	0	44	0	0.0	0.0	984.0

$${}_{25}q_{20}^A = 1 - (984/1000) = 0{,}01600$$

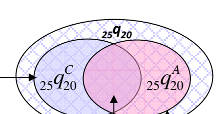
$$P_x^R = P_x + 0{,}5 \cdot ({}_5e_x - {}_5s_x - \boxed{{}_5d_x^C})$$

la population au risque est ajustée aux décès en compétitions

$${}_5q_x^A = \frac{{}_5d_x^A}{P_x + 0{,}5 \cdot ({}_5e_x - {}_5s_x - {}_5d_x^C)}$$

$${}_5d_x^A = S_x^A \cdot {}_5q_x^A$$

$$S_{x+5}^A = S_x^A - {}_5d_x^A$$



La probabilité complète : $${}_{25}q_{20} = {}_{25}q_{20}^C + {}_{25}q_{20}^A - {}_{25}q_{20}^C \cdot {}_{25}q_{20}^A = 0{,}1717$$

Table associée à un seul facteur dans un processus d'extinction multiple

Les questions qui se posent sont suivant :

1. Quelle sera la table de mortalité s'il n'existe qu'*une seule et unique cause* de décès? (la table nette de la mortalité d'une cause, ou la table associée à une cause de décès)

2. Quelle sera *la table de mortalité en absence total d'une des causes* de décès? ('*cause-deleted table'/ « table cause éliminée »*)

Table associée à un seul facteur de décroissance : les bases mathématiques

Dans une processus d'extinction multiple chaque facteur de décroissance « i » est associé à une fonction de la force de décroissance $\mu^i(x)$, qui est $\mu^i(x) = \lim_{n \to 0} {}_n m_x^i$

Comme les causes de sortie sont indépendantes, ➔ ${}_n m_x^1 + {}_n m_x^2 + \ldots + {}_n m_x^k = {}_n m_x$

➔ et par conséquent ${}_n\mu_x^1 + {}_n\mu_x^2 + \ldots + {}_n\mu_x^k = {}_n\mu_x$

On sait (*cours 1, diapositive 11*) que ➔ ${}_n p_x = e^{-\int_x^{x+n} \mu(y)dy}$

$$\boxed{{}_n p_x = e^{-\int_x^{x+n}[\mu_y^1 + \mu_y^2 + \ldots + \mu_y^k]dy} = e^{-\int_x^{x+n}\mu^1(y)dy} \cdot e^{-\int_x^{x+n}\mu^2(y)dy} \cdot \ldots \cdot e^{-\int_x^{x+n}\mu^i(y)dy} \cdot \ldots \cdot e^{-\int_x^{x+n}\mu^k(y)dy}}$$

ou ${}_n p_x = {}_n\overset{*}{p}{}_x^1 \cdot {}_n\overset{*}{p}{}_x^2 \cdot \ldots \cdot {}_n\overset{*}{p}{}_x^2$ et où ➔ $\boxed{{}_n\overset{*}{p}{}_x^i = e^{-\int_x^{x+n}\mu^i(y)dy}}$ (1)

« * » signifie une fonction relative à extinction unique

Construction des tables associées à une seule cause de sortie

L'expression (1) représente la probabilité de survivre dans les conditions de présence d'une seule cause (i) de sortie de l'état (un seul facteur de décroissance).

$$\left.\begin{aligned} {}_n d_x &= S(x)\cdot \int_x^{x+n} e^{-\int_x^a \mu(y)dy}\mu(a)da \\ {}_n^* d_x^i &= {}^*S^i(x)\cdot \int_x^{x+n} e^{-\int_x^a \mu^i(y)dy}\mu^i(a)da \\ {}_n d_x^i &= S(x)\cdot \int_x^{x+n} e^{-\int_x^a \mu^i(y)dy}\mu^i(a)da \end{aligned}\right\}$$

Par conséquent on peut appliquer une algorithme classique de construction d'une table de mortalité à partir des taux ou des quotients

Cependant, on ne peut aucunement observer directement $\mu^i(x)$ ni ${}_n^* m_x^i$

Méthode 1 :

On néglige une différence possible entre ${}_n^* m_x^i$ et ${}_n m_x^i$

Hypothèse : taux de mortalité à cause de *i* restera inchangé, si toutes autres causes de décès sont éliminées (voir le diapositive 8 sur l'indépendance des taux et la dépendance de la probabilité)

Dans ce cas ${}_n^* m_x^i = {}_n M_x^i = {}_n m_x^i$ et puisque ${}_n m_x^{-i} = 0$

on obtient $${}_n^* q_x^i = \frac{n\cdot {}_n m_x^i}{1+(n-{}_n^* a_x^i)\,{}_n m_x^i}$$ (voir F2, sur le diapositive 7)

Il reste un problème d'estimation correcte de ${}_n^* a_x^i$

Plus la force de mortalité est élevée, moins la vie moyenne est longue : dans le cas d'une seule cause de décès, il est difficile d'imaginer que l'âge moyenne au décès est indifférente à la durée d'intervalle, quoique une telle hypothèse soit acceptable dans le cas de la multitude des causes et de leur compensation mutuelle…

Méthode 2

Si la force de la décroissance associée au facteur *i* est *constante* (et intégrable) sur l'intervalle x, x+n,

alors $${}_n^{*}p_x^i = e^{-\int_x^{x+n} {}_n m_x^i \, dx} = e^{-n \cdot {}_n M_x^i}$$

et $${}_n^{*}L_x^i = \frac{{}^{*}S_x^i - {}^{*}S_{x+n}^i}{{}_n M_x^i}$$

on n'a pas donc besoin de « a_x » pour calculer la durée de vie

C'est une méthode est logique et simple, mais les résultats ne sont bons que si l'intervalle *n* est suffisamment court….

Méthode 3 (de C. Chiang, 1968)

Chiang, C.L. (1968) *An Introduction to Stochastic Processes in Biostatistic*. NY: Wiley

Soit risques associés à une cause *i* sont proportionnelles au risque total :

$$\mu^i(a) = R^i \cdot \mu(a) \qquad \text{pour } x \le a \le x+n$$

Alors on peut transformer l'équation (1, p.18): $${}_n^{*}p_x^i = e^{-\int_x^{x+n} \mu^i(a)da} = e^{-\int_x^{x+n} R^i \cdot \mu(a)da}$$

et puisque R^i est une constante : → $${}_n^{*}p_x^i = e^{-R^i \cdot \int_x^{x+n} \mu^i(y)dy} = \left[e^{-\int_x^{x+n} \mu^i(y)dy} \right]^{R^i} \Rightarrow$$

$$\boxed{{}_n^{*}p_x^i = \left[{}_n p_x\right]^{R^i}} \qquad (2)$$

Il ne reste qu'estimer R^i ☺

Estimation de R^i

Soit *N(a)* est une population soumise au risque, alors :

$$\frac{{}_nD_x^i}{{}_nD_x} = \frac{\int_x^{x+n} N(a)\cdot R^i \cdot \mu(a)\, da}{\int_x^{x+n} N(a)\cdot \mu(a)\, da} = R^i$$

et la substitution dans l'équation (2) donne : $\boxed{{}_n^*p_x^i = {}_np_x^{\left(\frac{{}_nD_x^i}{{}_nD_x}\right)}}$

Pour construire une table de mortalité avec « une ou plusieurs causes de décès éliminées », on remplace R^i par R^{-i} :

$${}_n^*p_x^{-i} = {}_np_x^{\,R^{-i}} = {}_np_x^{\left(\frac{{}_nD_x - {}_nD_x^i}{{}_nD_x}\right)}$$

Méthodes de l'estimations de « ${}_na_x^i$ »

1. Hypothèse de l'équivalence pour toute les causes *i*,

   ➔ ${}_n^*a_x^i = {}_na_x$

   il est cependant évident qu'une table avec une force de mortalité $\mu^i(a)$ plus élevée a une distribution des années vécues ***plus jeune*** et par conséquent les valeurs de ${}_n^*a_x^i$ plus petites

2. Calibrage de ${}_n^*a_x^i$ (N.Keifitz, 1966, « A Life table that argees with the Data » *Journal of American Statistical Association,* 63 (324): 1252-68)

$${}_n^*a_x^i = \frac{-\frac{n}{24}\cdot{}_n^*d_{x-5}^i + \frac{n}{2}\cdot{}_n^*d_x^i + \frac{n}{24}\cdot{}_n^*d_{x+5}^i}{{}_n^*d_x^i} \qquad (3)$$

+ on calcule ${}_n^*d_x^i$ directement sans recours à ${}_n^*a_x^i$

– ne convient pas aux intervalles initial et final

Méthodes de l'estimations de « ${}_na_x^i$ » (suite)

3. Interpolation entre deux situations extrêmes :

1 – il n'y pas de décès à cause de *i* dans l'intervalle *x,x+n* → alors la durée de vie dans l'intervalle est égale à *n*

2 – tous les décès dans l'intervalle sont causés par *i*, → alors ${}_n^*a_x^i = {}_na_x$

Respectivement dans ces situations extrême soit $R^i = 0$, soit $R^i = 1$

Pour les cas intermédiaires $\frac{{}_n^*L_x^i}{{}^*S^i(x)} = n - R^i\left(n - \frac{{}_nL_x}{S(x)}\right)$ = (la durée d'intervalle) – (la part des années perdues, attribué à la cause i) (4)

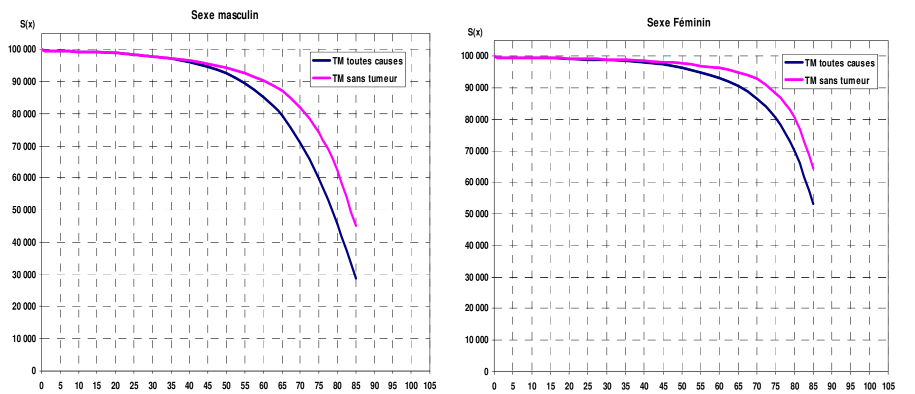
Puisque dans toutes les tables de mortalité: $\frac{{}_nL_x}{S(x)} = n\cdot{}_np_x + {}_na_x\cdot{}_nq_x = n-(n-{}_na_x)\cdot{}_nq_x$

On peut donc transformer l'équation (4) ${}_n^*a_x^i = n + R^i\frac{{}_nq_x}{{}_n^*q_x^i}({}_na_x - n)$ (5)

Pour le dernier intervalle: ${}_\infty^*a_x^i = {}^*e_\omega^i = \frac{e_\omega^0}{R^i}$ (6)

Exemple d'une table associée à une seule cause d'extinction (une cause éliminée)

Table de mortalité « toutes causes confondue » et table « cause tumeur éliminée » (France, 1999)



Les mêmes données qu'on a utilisées pour construire une table de mortalité d'extinction multiple. (voir le chapitre précédent et l'exercice)

Ces graphiques ne sont pas beaux à cause de la troncature des données au-delà de l'âge de 85 ans

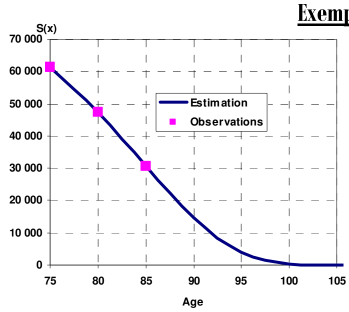
Application de la loi de Gompertz[1] pour « clôturer » (extrapoler) les table la mortalité

$$\mu(x) = \alpha \cdot e^{\beta \cdot x} \Rightarrow S(x) = C \cdot a^{b^x}$$

On peut estimer les paramètres C, a et b à partir de trois valeurs de la table de mortalité.

$$b = \left[\frac{\ln \frac{S(x+2n)}{S(x+n)}}{\ln \frac{S(x+n)}{S(x)}} \right]^{\frac{1}{n}} ; \quad a = \exp\left(\frac{\ln \frac{S(x+n)}{S(x)}}{b^x \cdot (b^n - 1)} \right); \quad C = S(x) \cdot \exp(-b^x \cdot \ln a)$$

Exemple: *France, table de mortalité 2000-2002*



$$b = \left[\frac{\ln \frac{S(85)}{S(80)}}{\ln \frac{S(80)}{S(75)}} \right]^{\frac{1}{5}} = 1{,}113403$$

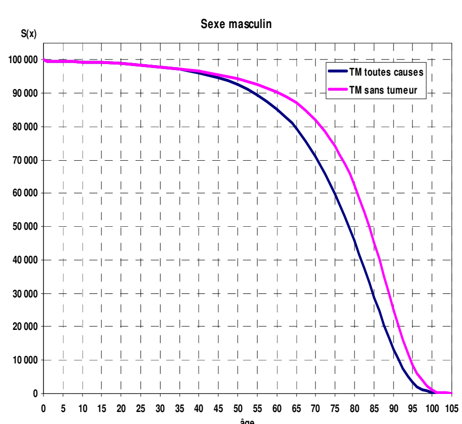
$$a = \exp\left(\frac{\ln \frac{S(80)}{S(75)}}{b^{75} \cdot (b^5 - 1)} \right) = 0{,}999886$$

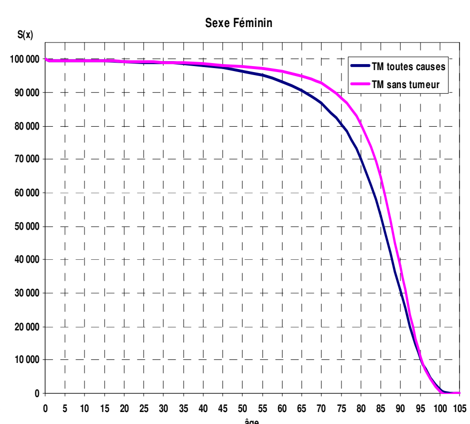
$$C = S(75) \cdot \exp(-b^{75} \cdot \ln a) = 87860{,}06$$

x	S(x)	Ŝ(x)
75	61250	*61250*
80	47391	*47391*
85	30554	*30554*
90		*14418*
95		*3988*
100		*442*
105		*10*
110		*0*

1) B.Gompetz (1825) – "On the nature of the Function expressive of the Law of human Mortality, and a new mode of determining the value of life contingencies" *Philosophical Transactions*, 115, p.513-583

Table de mortalité cause tumeur éliminée (France, 1999)





La survie est extrapolée au-delà de l'âge 85 ans selon la loi de Gompertz

Décomposition d'une différence entre les deux espérances de vie avec la prise en considération des causes de décès

1. On fait d'abord une décomposition sans prendre en considération des causes de décès selon la méthode d'Arriaga :

$$ {}_n\Delta_x = \frac{S_x^1}{S_0} \cdot \left(\frac{{}_nL_x^2}{S_x^2} - \frac{{}_nL_x^1}{S_x^1} \right) + \frac{T_{x+n}^2}{S_0} \cdot \left(\frac{S_x^1}{S_x^2} - \frac{S_{x+n}^1}{S_{x+n}^2} \right) $$

2. Ensuite, on doit ajuster ${}_n\Delta_x$ pour chaque intervalle d'âge proportionnellement des différences entre les risques attribués à une cause *i* :

$$ {}_n\Delta^i{}_x = {}_n\Delta_x \cdot \frac{{}_nm_x^i(2) - {}_nm_x^i(1)}{{}_nm_x(2) - {}_nm_x(1)} = $$

$$ = {}_n\Delta^i{}_x = {}_n\Delta_x \cdot \frac{{}_nR_x^i(2) \cdot {}_nm_x(2) - {}_nR_x^i(1) \cdot {}_nm_x(2)}{{}_nm_x(2) - {}_nm_x(1)} $$