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

Chapitre 6 :

Analyse des interférences : tables d'extinction multiple

1. Événements 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 tables combinées pour analyser des événements concurrents
5. Construction des tables épurées « nettes » (supposant l'absence des événements concurrents)
6. Tables associées à une seule cause de sortie (de décès)
7. Table de mortalité comme un modèle de survie en continu
8. Analyse de variance des indicateurs d'une table de mortalité
   (si ce sujet n'a pas été abordé dans le chapitre 5)

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)

(sans précaution) C. L. Chiang, *Introduction in stochastic Processes in Biostatistics*, 1968, Ch.9 et 10

Les événements perturbateurs et concurrents

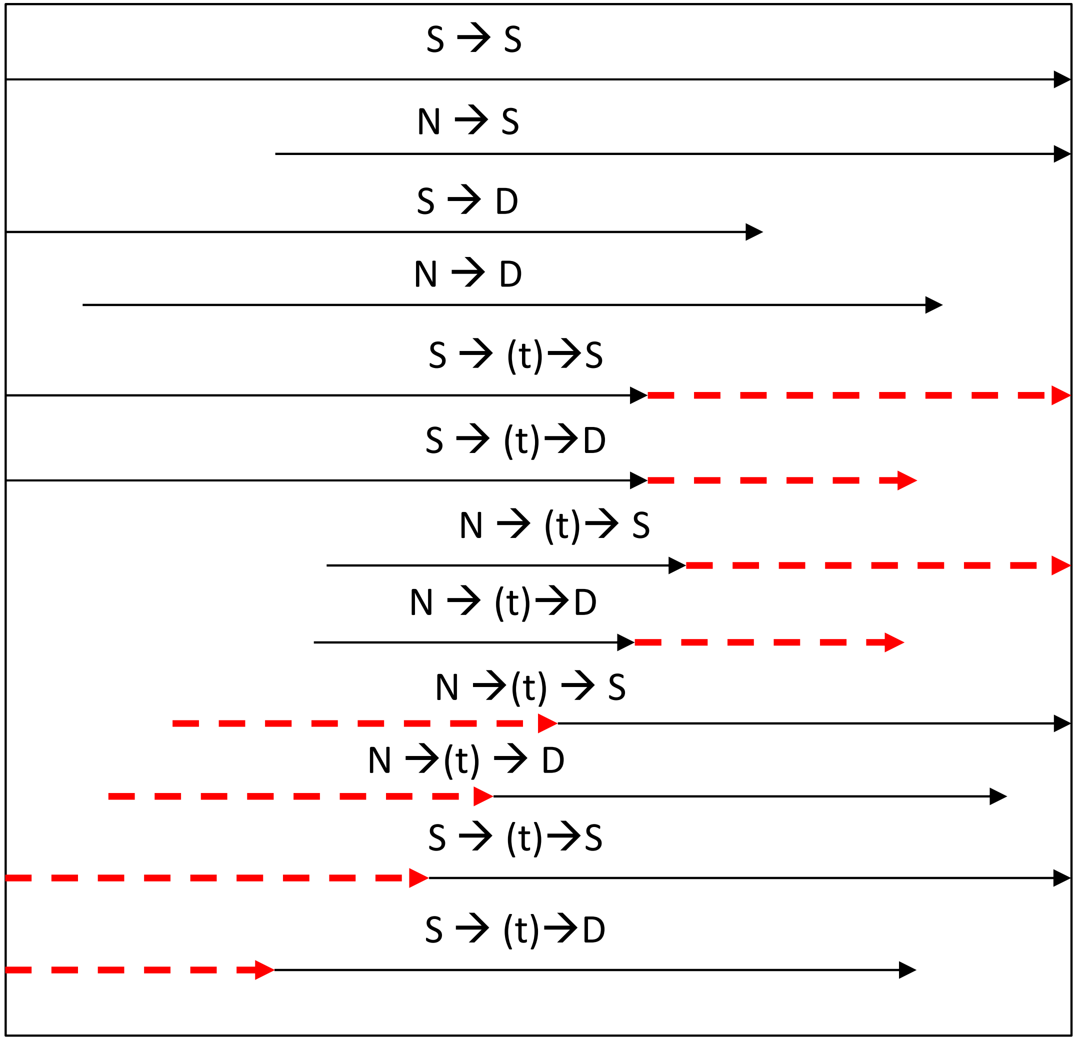
Les événements concurrents : l'arrivée E_1 empêche E_2 de se produire, et réciproquement (symétrie)

- Prendre un bus (celui qui arrive le premier)
- Décès par cause (on n'en meurt pas deux fois)
- Divorce et veuvage
- Sortie du célibat : mariage et décès

Les événements perturbateurs : l'arrivée de E_2 empêche E_1, mais la réciproque n'est pas vraie (variation de la population exposée)

- Si vous êtes en retard, vous ne pouvez pas prendre le bus
- Décès et émigration (le décès empêche l'émigration, l'émigration n'empêche pas le décès, quoi qu'il l'empêche se produire sur place)
- Mariage et émigration (l'émigration empêche le mariage sur place, mais mariage n'empêche pas d'émigrer)
- Divorce où veuvage empêche avoir une naissance dans le mariage…

Prise en considération de la migration dans le modèle générale de la population « démographique »



<u>États / évènements à considérer :</u>

S – survie

N – naissance (équivaut migration)

D – décès

t – transition d'une population à une autre

<u>**Quatre transitions** sont possibles sans perturbateur (migration spatiale ou sociale)</u>

S → S (un survivant survie)

S → D (un survivant décède)

N → S (un nouveau-né survie)

N → D (un nouveau-né décède)

<u>**Avec un phénomène perturbateur** (migration spatiale ou sociale) le nombre de transition possibles s'augmente **à douze**</u>

S → **(t)** →S (survivant part et survie ailleurs)

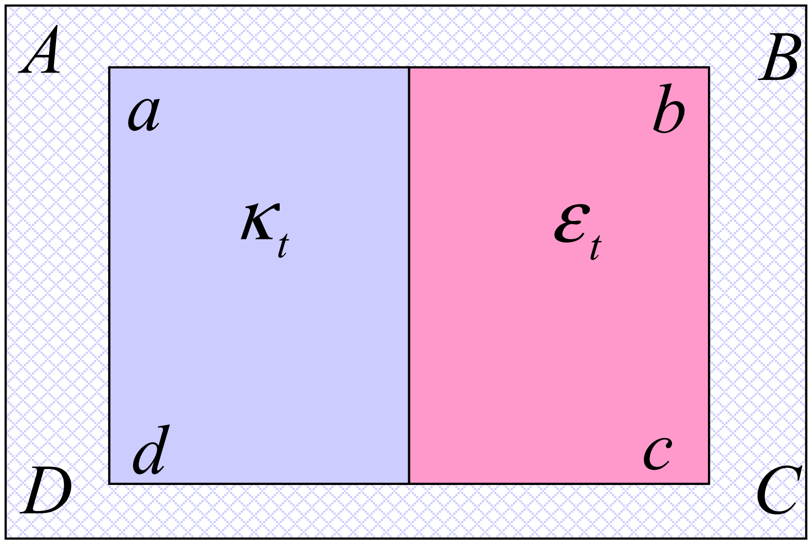
etc….

Les probabilités d'événements indépendants et les probabilités complémentaires (« brutes »)

Soit

$k_t = \dfrac{D_t}{P_t}$ – probabilité de mourir $\qquad \varepsilon_t = \dfrac{E_t}{P_t}$ – probabilité de partir

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



Soit

$\boldsymbol{S}$ la surface du rectangle *ABCD* représentant l'univers d'une expérience aléatoire (toutes les issues possibles)

$\boldsymbol{s}$ la surface du rectangle *abcd* représentant les issues (les événements) de sortir de l'observation en raison du décès (**κ**) ou du départ (**ε**) qui sont mutuellement exclusives (disjoints) $\kappa \cup \varepsilon \in s$;

$-\boldsymbol{s} = \boldsymbol{S} - \boldsymbol{s}$ surface (que l'on considère comme « *anti s* » réunissant l'ensemble des issues qui ne sont ni **κ**, ni **ε**

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

Les probabilités brutes sont additives

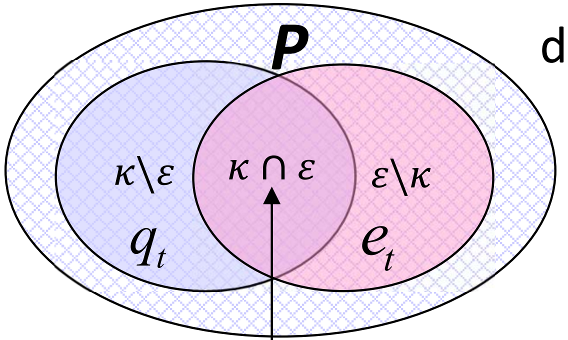
(cf. Les expériences aléatoires avec remise).

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

Les événements indépendants et concurrents (probabilités « nettes »)

Présentation de des probabilités nettes sur un diagramme de d'*Euler-Venn* (diagramme de Venn tout court) :

- 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 ***q*** (probabilité « nette » de mourir) et d'un cercle ***e*** *(probabilité « nette » de partir)*



Interférence entre κ et ε
$\kappa \cap \varepsilon \in s$;

« formule » de Berkson[1)]

$$q_t = \frac{D_t}{P_t - 0{,}5 \cdot E_i} \rightarrow$$ 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

Les probabilités nettes ne sont pas additives, mais leurs compléments sont conditionnels (cf. les expériences aléatoires sans remise) : $1 - s_t = (1 - q_t) \cdot (1 - e_t)$

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

donc $s_t = 1 - (1 - q_t) \cdot (1 - e_t) \rightarrow \boxed{s_t = q_t + e_t - q_t \cdot e_t}$

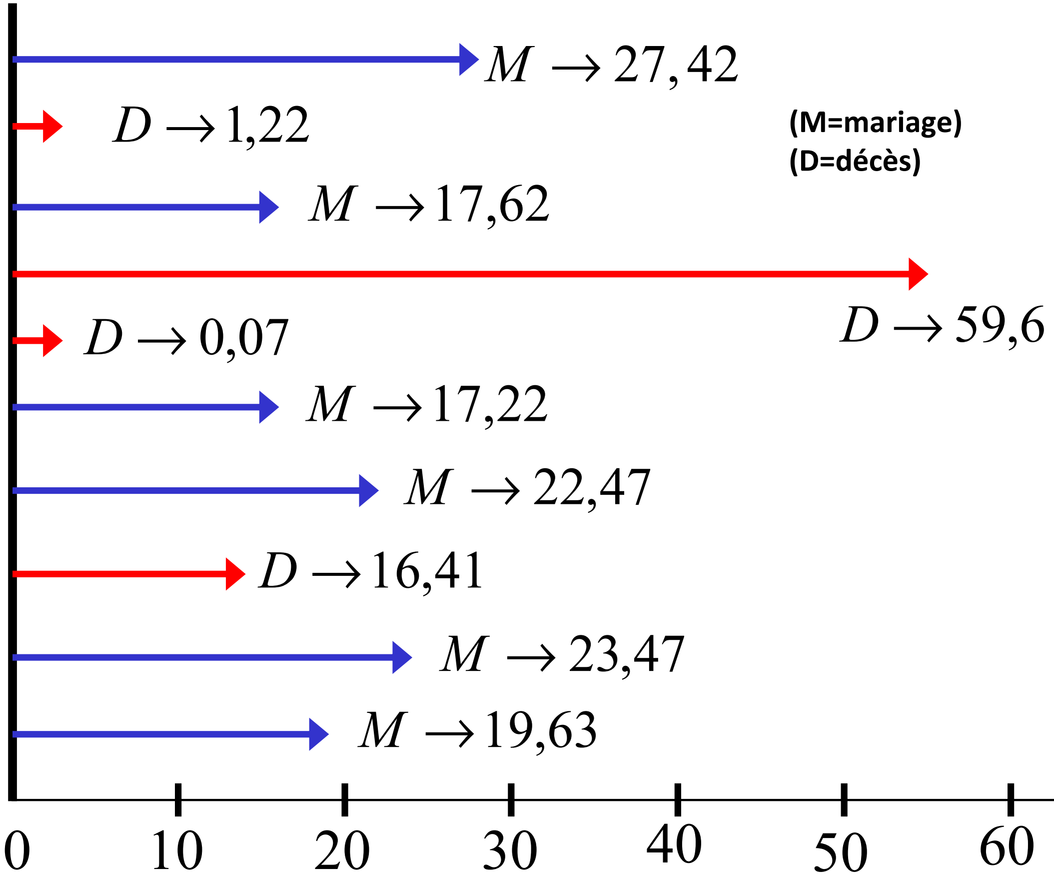
1) Berkson, Joseph (June 1946). "Limitations of the Application of Fourfold Table Analysis to Hospital Data". *Biometrics Bulletin*. 2 (3): 47–53. doi:10.2307/3002000. JSTOR 3002000.
D.Schwatz, Ph.Lazar (1964) "Taux de mortalité par une cause donnée de décès en tenant compte des autres causes de décès ou de disparition", *Révue de statistique appliquée* , tom 12,n°3 (1964) h.15-28 (http://www.numdam.org/item/?id=RSA_1964__12_3_15_0)

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 mesurée en absence des contributions des autres facteurs (expérience aléatoire avec remise).

- ***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 (expérience aléatoire sans remise)..
  On les appelle par ailleurs « les tables associées à un seul facteur dans le processus d'extinction multiple » ou tables associées à un seul facteur tout court.

Présentation des événements concurrents (données individuelles)*



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

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

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

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

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

* - cet exemple est emprunté aux Preston et al., 2001, p.73

Condition: tous les individus sont exposés au risque de décéder étant célibataire **et** au risque de se marier. Ces deux risques composent le risque total de sortie du célibat (risques combinés, additifs)

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

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

Une table combinée de nuptialité et de mortalité

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

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

Table combinée pour une période :

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

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

le taux ${}_n m_x^i = \frac{{}_n d_x^i}{{}_n L_x}$ et le quotient (probabilité) ${}_n q_x^i = \frac{{}_n d_x^i}{l_x}$ sur l'intervalle x, x+n

sachant que ${}_n d_x^i = {}_n m_x^i \cdot {}_n L_x$ et $l_x = \frac{1}{n} \cdot \left[{}_n L_x + (n - {}_n a_x) \cdot {}_n d_x \right]$

ou bien $\frac{{}_n L_x}{l_x} = {}_n p_x \cdot n + {}_n q_x \cdot {}_n a_x$; la conversion des taux en probabilité nous donne:

$$ {}_n q_x^i = \frac{n \cdot {}_n m_x^i}{1 + (n - {}_n a_x) \cdot {}_n m_x} \qquad \textbf{(F1)} $$

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

$$ {}_n m_x = {}_n m_x^i + {}_n m_x^{-i} \quad \Longrightarrow \quad {}_n q_x^i = \frac{n \cdot {}_n m_x^i}{1 + (n - {}_n a_x) \cdot ({}_n m_x^i + {}_n m_x^{-i})} \qquad \textbf{(F2)} $$

« $-\boldsymbol{i}$ » = tous événements excepté l'événement « i »

Relation entre les taux et les probabilités partiels dans une table (dans un modèle) de risques combinés

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

si ${}_nm_x^i = \dfrac{{}_nd_x^i}{{}_nL_x}$ est constant alors plus grand est de $\boldsymbol{{}_nm_x^{-i}} \rightarrow$ plus petit est $\boldsymbol{{}_nq_x^i}$

On considère donc $\boldsymbol{{}_nq_x^i}$ comme ***la probabilité dépendante***

En revanche, la diminution de $\boldsymbol{{}_nm_x^{-i}}$ n'a que peu, voire aucune influence sur $\boldsymbol{{}_nm_x^i}$ puisque 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 n'augmente pas forcément, si le taux de mortalité par autres cause diminue

Calculs d'une table combinée d'extinction multiple

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

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

Si nous acceptons une hypothèse que: ${}_nm_x^i = {}_nM_x^i \quad et \quad {}_nm_x = {}_nM_x$

où ***M*** – taux observé

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

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

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

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

3) $$l_x^i = \sum_{a=x}^{\infty} {}_nd_x^i$$

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

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

Age exact	Décès total	Décès de cancer	l_x	${}_nq_x$	${}_nq^i_x$	${}_nd^i_x$	l^i_x
x	${}_nD_x$	${}_nD^i_x$					
0	1 841	22	100 000	0.00494	0.00006	6	32 504
1	403	73	99 506	0.00110	0.00020	20	32 498
5	266	48	99 397	0.00071	0.00013	13	32 478
10	349	31	99 326	0.00089	0.00008	8	32 465
15	1 446	130	99 238	0.00358	0.00032	32	32 457
20	2 126	159	98 883	0.00566	0.00042	42	32 425
25	2 419	181	98 323	0.00574	0.00043	42	32 383
30	2 806	281	97 759	0.00658	0.00066	64	32 341
35	3 981	796	97 116	0.00925	0.00185	180	32 276
40	6 506	1 952	96 218	0.01545	0.00464	446	32 096
45	9 944	4 176	94 731	0.02337	0.00982	930	31 650
50	12 940	6 276	92 517	0.03355	0.01627	1505	30 720
55	12 934	6 273	89 413	0.04683	0.02271	2031	29 215
60	18 688	9 344	85 226	0.06917	0.03458	2948	27 184
65	27 814	12 516	79 331	0.10548	0.04747	3766	24 237
70	35 939	15 094	70 963	0.15774	0.06625	4701	20 471
75	43 859	15 351	59 769	0.23979	0.08393	5016	15 770
80	27 999	8 400	45 437	0.36668	0.11001	4998	10 754
85	62 504	12 501	28 776	1.00000	0.20000	5755	5 755
Total	274 764	93 605				**32 504**	

Table de la mortalité générale

Part de la mortalité générale associée aux tumeurs

1) $${}_nq^i_x = {}_nq_x \cdot \frac{{}_nD^i_x}{{}_nD_x}$$

2) $${}_nd^i_x = {}_nq^i_x \cdot l_x$$

3) $$l^i_x = \sum_{a=x}^{\infty} {}_nd^i_a$$

Proportion de nouveau-nés qui mourront éventuellement de tumeur:
32 504/100 000=32,5%

Proportion d'homme survécus jusqu'à l'âge de 75 ans qui mourront éventuellement de tumeur:
15 770/59 769=26,4%

Pour savoir + sur la mortalité en France :

MORTALITÉ

En raison d'un changement de méthode de codage d'enregistrement des cas, les données de mortalité actualisées ne sont pas disponibles.

RÉPUBLIQUE FRANÇAISE

INSTITUT NATIONAL DU CANCER

et c'est tout ce que l'INC peut nous communiquer en 2024

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

Cette analyse est basée sur une étude suivie (*follow up study*) des grossesses sur l'île Kauai pendant 4 ans de 1953 à 1956

- Facteur de risque : la durée de grossesse
- Événement d'intérêt : un décès intra-utérin (fausse couche)
- Événement concurrent : une naissance vivante

  NB : dans le cas de cette étude il est difficile à imaginer une situation dans laquelle risque d'un événement 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 femme avec une durée de grossesse quelconque 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)

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

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

Événement 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	Effectifs exposés au risque pour calculer les quotients	Quotients		Table de mortalité		
							Naissances vivantes	Décès	Survivants	Décès	Nais. vivantes
t	**Et**	**Dt**	**St**	**Nt**	(Et-Dt-St-Nt)= **G_{t+1}**	Gt+(Et-St)/2= **P_R**	**v(t)**	**q(t)**	**S(t)**	**d(t)**	**n(t)**
4-8	592	32	0	0	0	0+(592-0)/2=296	0	0.1081	1 000.0	108.11	0
8-12	941	72	1	0	592-32-0-0=560	1030	0	0.0699	891.9	62.346	0
12-16	585	77	2	0	1428	1719.5	0	0.0448	829.5	37.147	0
16-20	337	28	2	0	1934	2101.5	0	0.0133	792.4	10.558	0
20-24	248	20	9	1	2241	2360.5	0.0004	0.0085	781.8	6.6244	0.3312
24-28	175	8	6	4	2459	2543.5	0.0016	0.0031	774.9	2.4372	1.2186
28-32	98	8	4	25	2616	2663	0.0094	0.003	771.2	2.3169	7.2402
32-36	67	8	6	72	2677	2707.5	0.0266	0.003	761.7	2.2506	20.255
36-40	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)→ sous-estimation de la probabilité de fausse-couche de 2.6 fois						8.95%	91.05%

Cet exemple a été utilisé dans Leridon, Toulemon, 1997, p.108-109

Analyse originale F.E.French, J.E.Bierman (1962) « Probablility of fœtal mortality » *Public Health Reports*, 77 (10) p.835-847

données tirées de la table 5 sur la page 841, les résultats de l'analyse de survie (table de mortalité) dans la table 6, page 842 (une copie est disponible sur l'EPI)

Table associée à un seul facteur de risque dans un processus d’extinction multiple (« table nette ») :

un exemple d’étude sur la mortalité de pilotes de Formule 1

- Facteur de risque : l’âge[1)]
- Événements d’intérêt : un décès (en et hors compétition)
- Événements concurrents : décès en compétitions et décès hors compétition

NB : dans le cas de cette étude 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 de décès net associé à la 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 quitté la carrière 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*** → proc lifereg; proc lifetest, ***SPSS*** et ***Statsoft*** → survival analysis; ***STATA*** → ltable; stcox, etc.)

Table « nette » : mortalité des pilotes

Exemple emprunté de Leridon, Toulemon, 1997, p.103-107

Analyse originale cf. J.-L.Rallu, (1991) « Sélection, âge et performance des coureurs de formule 1 » *Population*, 46,6, p.1711-1733

1. 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(-24)	10	0	0	0	0
25(-29)	100	20	4	4	0
30(-34)	206	18	9	8	1
35(-39)	10	40	7	6	1
40(-44)	0	136	4	3	1
45(+)	0	88	0	0	0

Condition d'analyse :

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

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

Hypothèse 1 :

la densité de risque (hasard) est uniforme à l'intérieur de chaque intervalle d'âge

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) :

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

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

$${}_5d_x = l_x \cdot {}_5q_x \qquad l_{x+5} = l_x - {}_5d_x$$

2. Table de mortalité générale (toutes causes confondues)

x	P_x	P_{x+5}	*Population pour calculer* ${}_nq_x$ P_x^R	${}_5D_x$	${}_5q_x$ *x1000*	${}_5d_x$	l_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	(86+206/2-18/2)=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 carrière d'un pilote de formule 1 est :

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

suite

3. Table de mortalité associée aux décès en compétition

x	P_x	P_{x+5}	Population initiale estimée (P_x^R)	${}_nD_x^C$	${}_nq_x^C$	${}_nd_x^C$	l_x^C
20	0	10	5	0	0.0	0.0	1000
25	10	86	(10+100/2-20/2-0/2)=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 à risque est ajustée aux décès hors compétitions (qui sont exclus)

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

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

$$l_{x+5}^C = l_x^C - {}_5d_x^C$$

4. Table de mortalité associée aux décès hors compétition

x	P_x	P_{x+5}	P^R	${}_nD_x^A$	${}_nq_x^A$	${}_nd_x^A$	l_x^A
20	0	10	5	0	0.0	0.0	1000
25	10	86	48	0	0.0	0.0	1000
30	86	266	176	1	5.7	5.7	1000
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

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

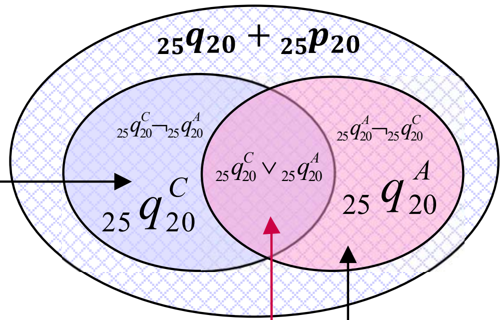
$$P_x^R = P_x + 0{,}5\cdot({}_5e_x - {}_5s_x - \boxed{{}_5d_x^C})$$

la population au risque est ajustée aux décès en compétitions

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

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

$$l_{x+5}^A = l_x^A - {}_5d_x^A$$



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

Disjonction (logique)

Pour passer de deux tables associées à une table combinée, il faudra distribuer la probabilité associée à la disjonction $C \vee A$ entre ${}_{25}q_{20}^C$ et ${}_{25}q_{20}^A$ soit en parties égales, soit proportionnellement aux nombre d'événements C et A

Table de mortalité avec les relations dans les termes de la continuité :
passage d'un intervalle défini [*x, x+n)* au voisinage d'un point x.

+ on n'a pas besoin d'hypothèse sur la distribution des décès ou de la force de mortalité sur les intervalles d'âge (*x, x+n*) puisque n → 0 par conséquent :

(1) $S(x) = S(a) \cdot e^{-\int\limits_a^x \mu(y)dy}$ pour x > a ➔ ${}_n p_x = \frac{l_{x+n}}{l_x} = \frac{S(x) \cdot e^{-\int\limits_x^{x+n} \mu(a)da}}{S(x)} = e^{-\int\limits_x^{x+n} \mu(a)da}$

(2) $${}_n p_x = e^{-\int\limits_x^{x+n} \mu(a)da}$$

(3) $${}_n d_x = \int\limits_x^{x+n} S(a) \cdot \mu(a)da = \int\limits_x^{x+n} S(x) \cdot e^{-\int\limits_x^a \mu(y)dy} \cdot \mu(a)da = S(x) \cdot \int\limits_x^{x+n} e^{-\int\limits_x^a \mu(y)dy} \cdot \mu(a)da$$

(4) $${}_n q_x = \frac{{}_n d_x}{l_x} = \frac{S(x) \cdot \int\limits_x^{x+n} e^{-\int\limits_x^a \mu(y)dy} \cdot \mu(a)da}{S(x)} = \int\limits_x^{x+n} e^{-\int\limits_x^a \mu(y)dy} \cdot \mu(a)da$$

(5) $$T_x = \int\limits_x^{\infty} S(a)da = S(x) \cdot \int\limits_x^{\infty} e^{-\int\limits_x^a \mu(y)dy} da$$

Dans la formule (1)
la limite inférieure de l' intégration = *a* .
Dans la formule (5)
la limite inférieure de l' intégration = *x* .

(6) $$e_x = \frac{T_x}{l_x} = \frac{\int\limits_x^{\infty} S(a)da}{S(x)} = \frac{S(x) \cdot \int\limits_x^{\infty} e^{-\int\limits_x^a \mu(y)dy} da}{S(x)} = \int\limits_x^{\infty} e^{-\int\limits_x^a \mu(y)dy} da$$

μ(x) – la fonction de la « force de mortalité » dans le point x ;
On voit que toutes les fonctions d'une TM sont définies par *μ(x)* et seulement par *μ(x)*

$$\mu(x) = \lim_{n \to 0} {}_n m_x$$

Taux de mortalité est un estimateur de *μ(y)* force de mortalité moyenne sur l'intervalle (x,x+n]

Pouvez-vous déduire la formule pour ${}_nL_x$ et ${}_nm_x$ sachant que

$${}_n m_x = \frac{{}_n d_x}{{}_n L_x}$$

Table (nette) 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 totale d'une des causes* de décès? ('*cause-deleted table'/ « table-une-cause-éliminée »*)

Definitions (cf. Chin L. CHIANG, *Introduction in stochastic Processes in Biostatistics*, 1968, p.242-243):

1. ***The crude probability***. The probability of death from a specific cause in presence of all other risks acting in population, or

   $Q_{i\delta}$=Pr{an individual alive at time x_i will die in the interval (x_i, x_{i+1}) from cause R_δ in presence of all other risks in the population} **(p.242/p.243)**

2. ***The net probability***. The probability of death if the specific risk is the only risk in effect in the population or, conversely, the probability of death if a specific risk is eliminated from the population.

   $q_{i.\delta}$= Pr{an individual alive at time x_i will die in the interval (x_i, x_{i+1}) if the cause R_δ is the only risk acting on the population}

   $q_{i.-\delta}$= Pr{an individual alive at time x_i will die in the interval (x_i, x_{i+1}) if R_δ is eliminated as a risk of death}

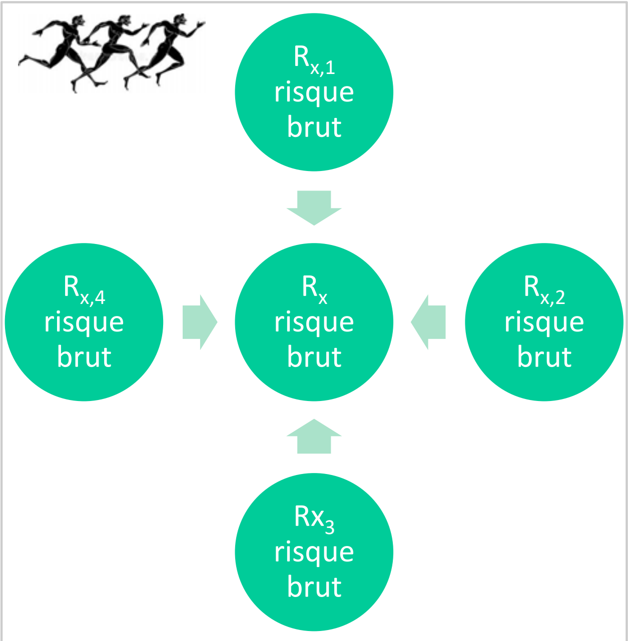
3. ***The partial crude probability***. The probability of death from a specific cause when another risk (or risks) is eliminated from the population.

   $Q_{i\delta.-1}$ = Pr{an individual alive at time x_i will die in the interval (x_i, x_{i+1}) from R_δ if R_1 is eliminated as a risk of death}

   $Q_{i\delta.-1,-2}$= Pr{an individual alive at time x_i will die in the interval (x_i, x_{i+1}) from R_δ if R_1 and R_2 is eliminated as a risk of death}

Notions de risque brut, net (épuré) et brut partiel (partiellement épuré)

Modèle 1 : tous les risques combinés

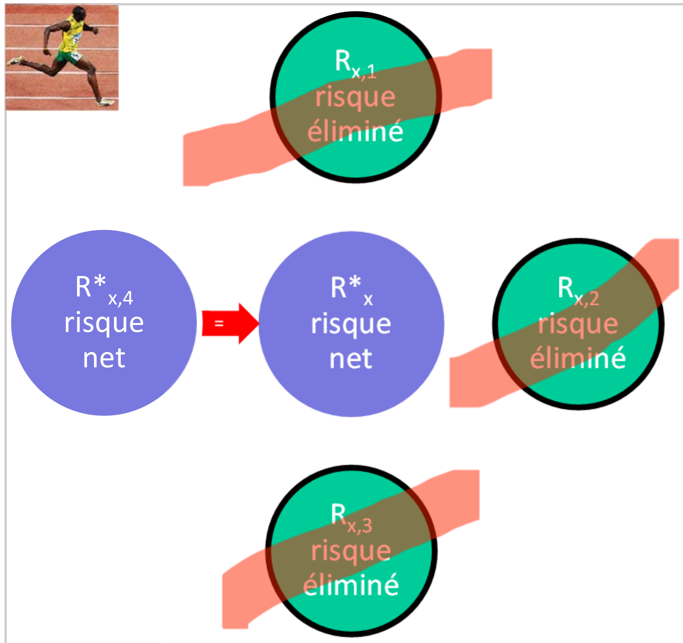


Tous les risques $R_{x,i}$ associés à des causes *i* (i=1; 2; …; n) et à l'âge *x* sont pris en compte.
Modèle de survie basé sur une table des risques combinés.

$$R_x = \sum_i R_{x,i}$$

Que le plus rapide gagne !

Modèle 2 : tous les risques éliminés sauf un

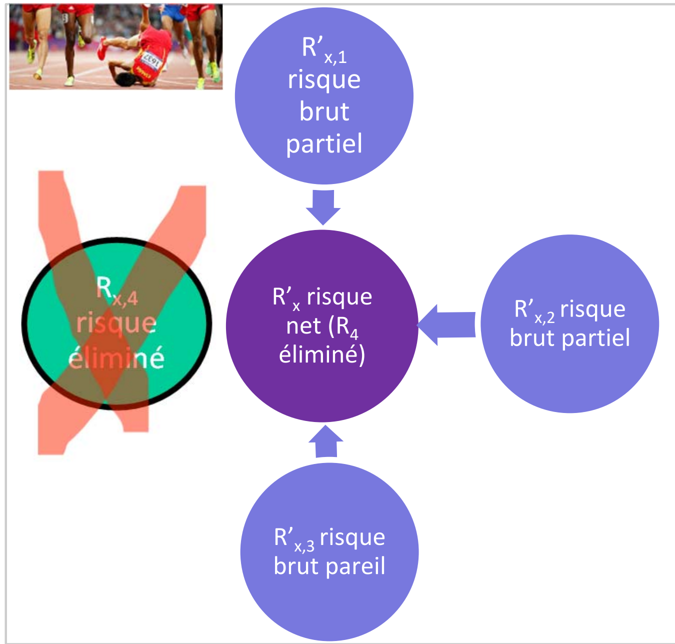


Un seul risque $R_{x,\delta}$ associé à une cause δ est pris en compte.
Modèle de survie basé sur une table de risque net associée à une seule cause de décès

$$R_x \neq R_x^* = R^*_{x,\delta} > R_{x,\delta}$$

Sans concurrence un seul est toujours gagnant, qu'il soit le plus rapide le plus lent ?

Modèle 3 : un seul risque éliminé



Un seul risque $R_{x,\delta}$ associé à une cause δ est éliminé tous les autres sont pris en compte.
Modèle de survie basé sur une table de risque net une cause éliminée

$$R_x \neq R_x^{'} = \sum_{i \neq \delta} R_{x,i}^{'}; \; R_{x,i}^{'} > R_{x,i}$$

Un est tombé, les autres ont plus de chance ! Sont-ils plus rapides ?

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

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

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

➔ par conséquent $\mu^1(x) + \mu^2(x) + \ldots + \mu^k(x) = \mu(x)$

Sachant que ${}_n p_x = e^{-\int_x^{x+n} \mu(y)dy}$ (*diapositive 18, formule 2*) on peut réécrire:

$$\boxed{{}_n p_x = e^{-\int_x^{x+n}\left[\mu_y^1 + \mu_y^2 + \ldots \mu_y^k\right]dy} = e^{-\int_x^{x+n} \mu_y^1 dy} \cdot e^{-\int_x^{x+n} \mu_y^2 dy} \cdot \ldots \cdot e^{-\int_x^{x+n} \mu_y^k dy}}$$

puisque probabilité de survie sont conditionnelles (i.e. pour survivre il faut ne pas mourir des causes 1, 2, …k) : ${}_n p_x = {}_x^* p_x^1 \cdot {}_x^* p_x^2 \cdot \ldots \cdot {}_x^* p_x^k$ où ${}_n^* p_x^i = e^{-\int_x^{x+n} \mu^i(y)dy}$ $\quad (1)$

« * » signifie une fonction relative à extinction unique, ou associée à une seule cause dans le processus d'extinction multiple

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).

(a) $${}_n d_x = \left[S(x)\cdot \int_x^{x+n} e^{-\int_x^a \mu(y)dy} \right]\cdot \mu(a)da$$

(b) $${}^*_x d^i_x = \left[{}^*S^i(x)\cdot \int_x^{x+n} e^{-\int_x^a \mu^i(y)dy} \right]\cdot \mu^i(a)da$$

(c) $${}_x d^i_x = \left[S(x)\cdot \int_x^{x+n} e^{-\int_x^a \mu^i(y)dy} \right]\cdot \mu^i(a)da$$

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

[l'expressions dans parenthèses rectangulaires sont les formules pour passer de la fonction continue $S(x)$ à des valeurs discrètes sur les intervalles $(x;\ x+n)$: $\boldsymbol{{}_nL_x}$ respectivement dans les tables de mortalité
a) toutes causes confondues;
b) associées à une cause unique dans le processus d'extinction multiple;
c) multiples causes combinées]

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

<u>Transcription des formules ***(il faut vivre pour mourir)*** :</u>

S(x) – le nombre de survivants à l'âge exact x = valeur de la fonction dans le point, multiplié par la probabilité de mourir sur l'intervalle x, x+n qui est une intégrale du produit de la fonction p(x) = probabilité de vivre à l'âge « a » et de la force de mortalité à l'âge « a » = µ(a) – cf. *diapositive 18, formule 4*

Méthode (approche) 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 10 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 $${}_{n}q_{x}^{i} = \frac{n \cdot {}_{n}m_{x}^{i}}{1 + (n - {}_{n}^{*}a_{x}^{i}) \cdot {}_{n}m_{x}^{i}}$$ (voir F2, sur ch. 6, diapositive 10)

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 moyen au décès est indifférent 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 (approche) 2

Si (c'est une hypothèse) la force de la décroissance associée au facteur *i* est <u>*constante*</u> (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 $$ {}_n^{*}L_x^i = \frac{{}^{*}l_x^i - {}^{*}l_{x+n}^i}{{}_n M_x^i} = \frac{{}_n^{*}d_x}{{}_n M_x^i} $$

pour calculer la durée de vie on n'a pas donc besoin de « $\boldsymbol{a_x}$ » (qui résulte de notre hypothèse)

C'est une méthode est logique et simple, mais les résultats ne sont bons que si l'intervalle *n* est suffisamment court, *et la plus souvent ce n'est pas le cas…*

Méthode 3 (risques proportionnels)

cf. Chin Long CHIANG (1968) *An Introduction to Stochastic Processes in Biostatistic*. NY : Wiley[1)]

Soit le risque associé à une cause *i* proportionnel au risque total : $\mu^i(a) = R^i \cdot \mu(a)$

on peut donc exprimer la probabilité de survie associée à un seul facteur d'extinction sur un intervalle $x < a \leq x+n$ dans un processus d'extinction multiple (cf. équation 1, diapositive 21) :

$$ {}_n^*p_x^i = e^{-\int_x^{x+n} \mu^i(a)da} = e^{-\int_x^{x+n} R^i \cdot \mu(a)da} $$

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

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

Il ne reste qu'estimer R^i ☺

1) Une démonstration un peu différente est présentée par Rolland Pressat (1995) *Eléments de la démographie mathématique*, p.83-88

Estimation du multiplicateur R^i

Soit $N(a)$ est une population exposée 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 : ${}_n^{*}p_x^i = {}_np_x^{\left(\frac{{}_nD_x^i}{{}_nD_x}\right)}$

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

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

Méthodes de l'estimations de « ${}_n a_x^{\,i}$ » dans une table de mortalité associée à la cause *i*

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

   ➔ ${}_n^{*}a_x^{\,i} = {}_n a_x$

   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)
   voir Chap.5 diapo. 31

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

\+ on calcule ${}_n^{*}d_x^{\,i}$ directement sans recours à ${}_n^{*}a_x^{\,i}$ (voir méthodes 2 et 3, diapos 24-25)

– ne convient pas aux intervalles initiaux et final

Suite : méthodes de l'estimations de « ${}_n a_x^i$ » pour les intervalles inégaux et terminaux

3. Interpolation entre deux situations extrêmes :

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

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

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

entre ces deux cas la durée de vie dans un intervalle d'une table ***associée*** est une moyenne entre ${}_n e_x$ et ***n*** pondérée par R^i $\Rightarrow \frac{{}_n^*L_x^i}{{}^*l_x^i} = n \cdot (1 - R_x^{\,i}) + R_x^{\,i} \cdot \frac{{}_n L_x}{l_x}$

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

Puisque dans toutes les tables de mortalité: $\frac{{}_n L_x}{l_x} = n \cdot {}_n p_x + {}_n a_x \cdot {}_n q_x = n - (n - {}_n a_x) \cdot {}_n q_x$

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

Pour le dernier (ω) intervalle : ${}_\infty^*a_\omega^i = {}^*e_\omega^i = \frac{e_\omega}{R_\infty^i}$ (6)

† cette démonstration est en annexe (à la fin de cette présentation)

Example: France, hommes 1999, tumeurs

$l_x,\ {}_np_x,\ {}_na_x, e_x$ = les fonctions de la table-maître ; $R_x^{-i} = \dfrac{{}_nD_x - {}_nD_x^i}{{}_nD_x}$; ${}_np_x^* = \left({}_np_x\right)^{R_x^{-i}}$; ${}^*l_{x+n}^{-i} = {}^*l_x^{-i} \cdot {}_n^*p_x^{-i}$

$${}_n^*a_x^{-i} = \frac{-\frac{n}{24}\cdot {}_n^*d_{x-5}^{-i} + \frac{n}{2}\cdot {}_n^*d_x^{-i} + \frac{n}{24}\cdot {}_n^*d_{x+5}^{-i}}{{}_n^*d_x^i}$$ pour x =5 à 75 $${}_n^*a_x^{-i} = n + R^{-i}\,\frac{{}_nq_x}{{}_n^*q_x^{-i}}\left({}_na_x - n\right)$$ pour x =0, 1, 80 $${}_\infty^*a_{85}^{-i} = {}^*e_{85}^{-i} = \frac{e_{85}^0}{R_{85}^{-i}}$$

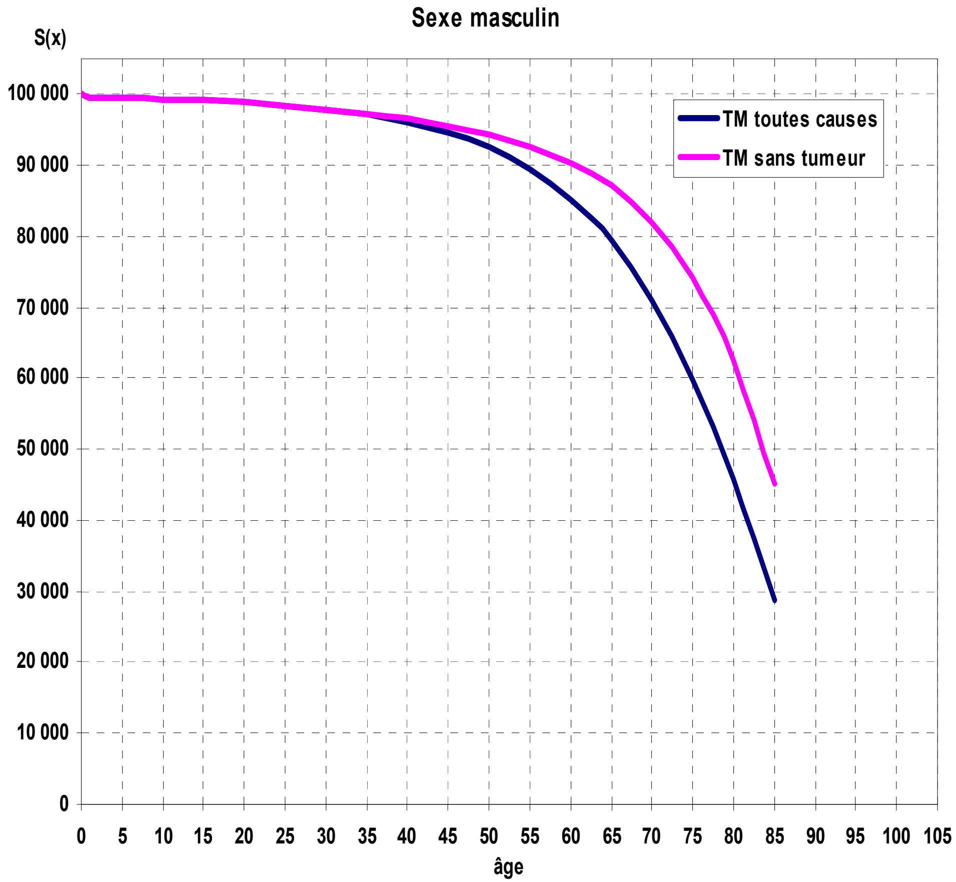
Table de mortalité-maître | **Table de mortalité associée à une cause éliminée**

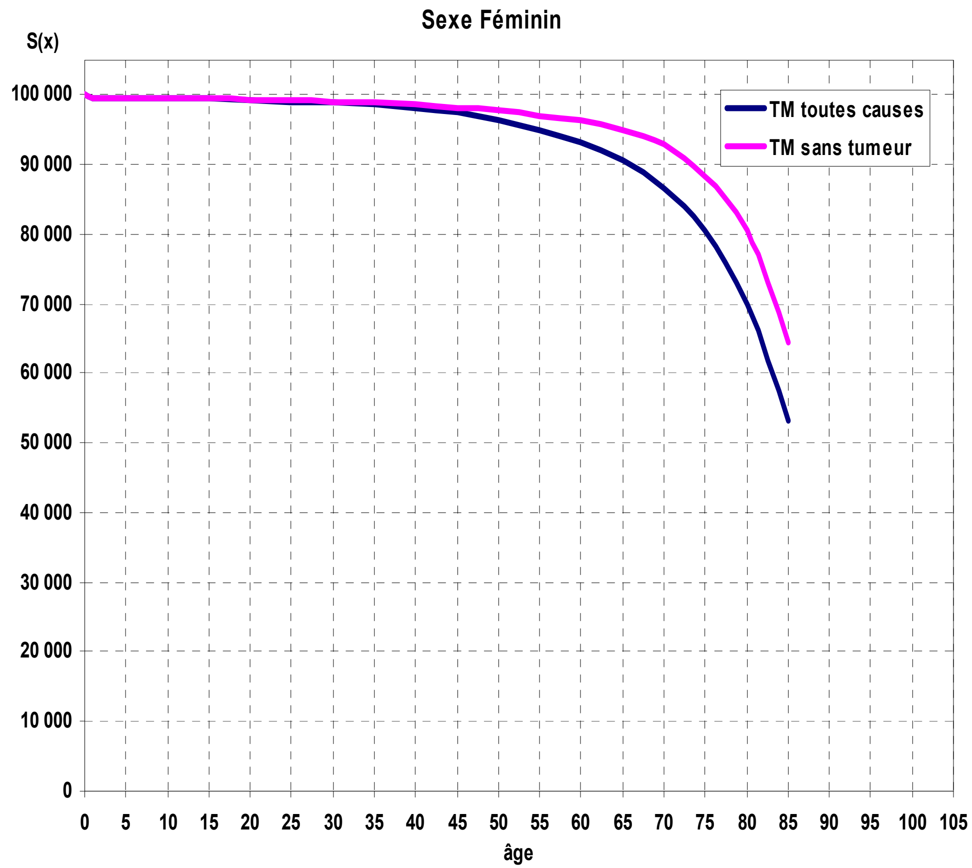
n	Age	l_x	R^{-i}	${}_np_x$	${}_na_x$	e_x	${}_n{}^*p_x$	${}^*l_x$	${}_n{}^*a_x$	${}_n{}^*L_x$	${}_n{}^*T_x$	${}^*e_x$
1	0	100 000	0.98800	0.9951	0.06000	74.98	0.99512	100 000	0.060028	99 541	7 997 751	79.978
4	1	99 506	0.82000	0.9989	1.60000	74.36	0.99910	99 512	1.600237	397 833	7 898 210	79.369
5	5	99 397	0.82000	0.9993	2.31000	70.44	0.99941	99 423	2.466803	496 909	7 500 377	75.439
5	10	99 326	0.91000	0.9991	2.86000	65.48	0.99919	99 364	3.189157	496 618	7 003 468	70.483
5	15	99 238	0.91000	0.9964	3.00000	60.54	0.99674	99 284	2.782548	495 835	6 506 850	65.538
5	20	98 883	0.92500	0.9943	2.50000	55.75	0.99476	98 961	2.580033	493 655	6 011 014	60.741
5	25	98 323	0.92500	0.9943	2.48000	51.05	0.99469	98 442	2.524454	490 948	5 517 359	56.047
5	30	97 759	0.90000	0.9934	2.57000	46.33	0.99408	97 920	2.571240	488 164	5 026 412	51.332
5	35	97 116	0.80000	0.9908	2.68000	41.62	0.99260	97 340	2.635233	484 950	4 538 247	46.623
5	40	96 218	0.70000	0.9845	2.72000	36.98	0.98916	96 619	2.615577	480 619	4 053 297	41.951
5	45	94 731	0.58000	0.9766	2.64000	32.52	0.98638	95 572	2.595153	474 754	3 572 678	37.382
5	50	92 517	0.51500	0.9664	2.62000	28.24	0.98258	94 270	2.621506	467 399	3 097 924	32.862
5	55	89 413	0.51500	0.9532	2.59000	24.13	0.97560	92 627	2.641882	457 762	2 630 525	28.399
5	60	85 226	0.50000	0.9308	2.63000	20.18	0.96480	90 368	2.691543	444 336	2 172 762	24.044
5	65	79 331	0.55000	0.8945	2.62000	16.49	0.94053	87 186	2.684445	423 963	1 728 426	19.825
5	70	70 963	0.58000	0.8423	2.59000	13.12	0.90523	82 002	2.685821	392 013	1 304 463	15.908
5	75	59 769	0.65000	0.7602	2.62000	10.10	0.83677	74 230	2.658652	343 111	912 451	12.292
5	80	45 437	0.70000	0.6333	2.59000	7.45	0.72633	62 114	2.739623	272 145	569 340	9.166
5.3	85	28 776	0.80000	0.0000	5.41000	5.27	0.00000	45 115	6.587500	297 195	297 195	6.588

Probabilité de survie à l'âge 85 ans toutes causes confondues = 28.7%; espérance de vie à la naissance =75 ans

Probabilité de survie à l'âge 85 ans les décès de tumeur éliminés = 45.1%; espérance de vie à la naissance = 80 ans

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

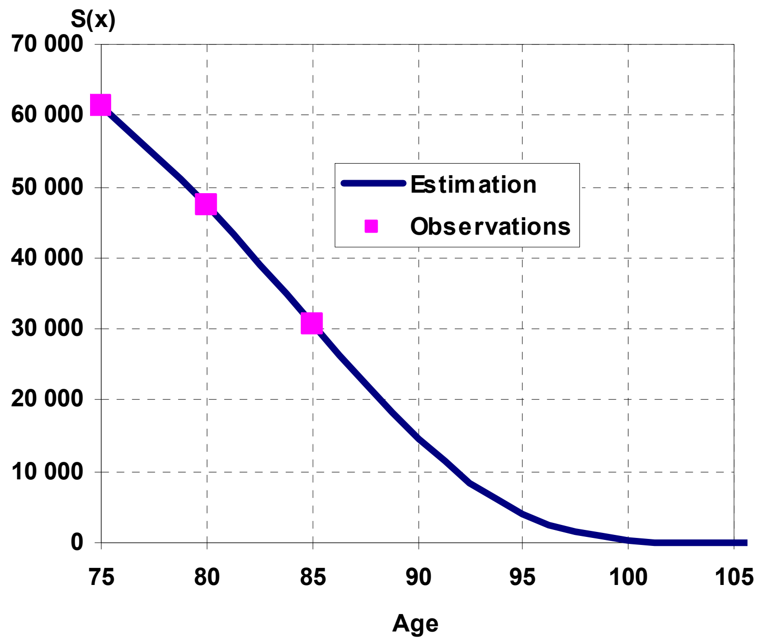
Application de la loi de Gompertz[1] pour « clôturer » (extrapoler) les table la mortalité

$\mu(x) = \alpha \cdot e^{\beta \cdot x}$ sachant que $\mu(x) = \frac{l(x)'}{l(x)}$ on peut en déduire $\Rightarrow$ $l_x = C \cdot a^{b^x}$

On estime ensuite les paramètres ***C***, ***a*** et ***b*** à partir de trois valeurs de la table de mortalité (système de trois équations avec trois valeurs inconnues).

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

<u>Exemple:</u> *France, table de mortalité 2000-2002*



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

$$a = \exp\left[\frac{\ln \frac{l_{80}}{l_{75}}}{b^{75} \cdot (b^5 - 1)} \right] = 0.999886$$

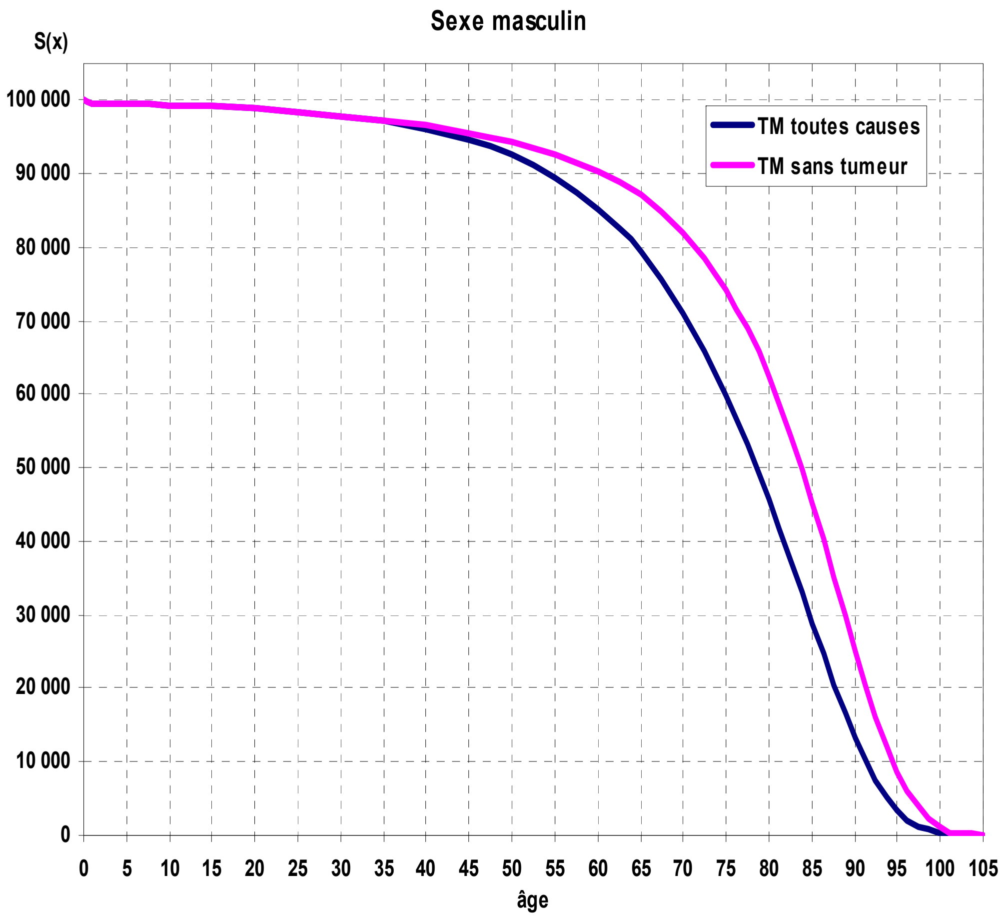
$$C = l_{75} \cdot \exp(-b^{75} \cdot \ln a) = 87860{,}06$$

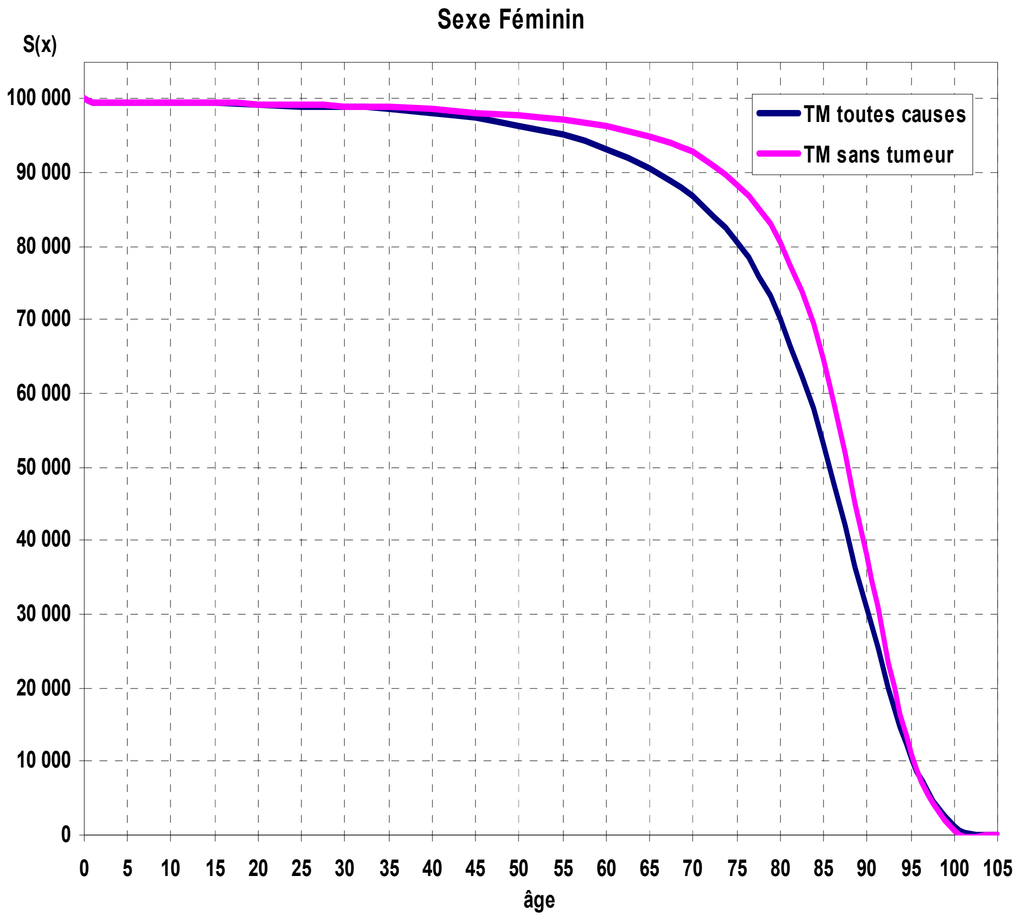
x	l_x	$\hat{l}_x$
75	61250	*61250*
80	47391	*47391*
85	30554	*30554*
90		*14418*
95		*3988*
100		*442*
105		*10*
110		*0*

Pour voir plus de détails cf.: Roland Pressat (1995) *Eléments de démographie mathématique*, AIDELF, p.11-14

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 une formule :

Andreev-Pressat :
$$ {}_n\Delta_x = \frac{0.5\cdot\left(l_x^1 + l_x^2\right)}{l_0}\cdot\left(e_x^2 - e_x^1\right) - \frac{0.5\cdot\left(l_{x+n}^1 + l_{x+n}^2\right)}{l_0}\cdot\left(e_{x+n}^2 - e_{x+n}^1\right) $$

Arriaga-Pollard :
$$ {}_n\Delta_x = \frac{l_x^1}{l_0}\cdot\left(\frac{{}_nL_x^2}{l_x^2} - \frac{{}_nL_x^1}{l_x^1}\right) + \frac{T_{x+n}^2}{l_0}\cdot\left(\frac{l_x^1}{l_x^2} - \frac{l_{x+n}^1}{l_{x+n}^2}\right) $$

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

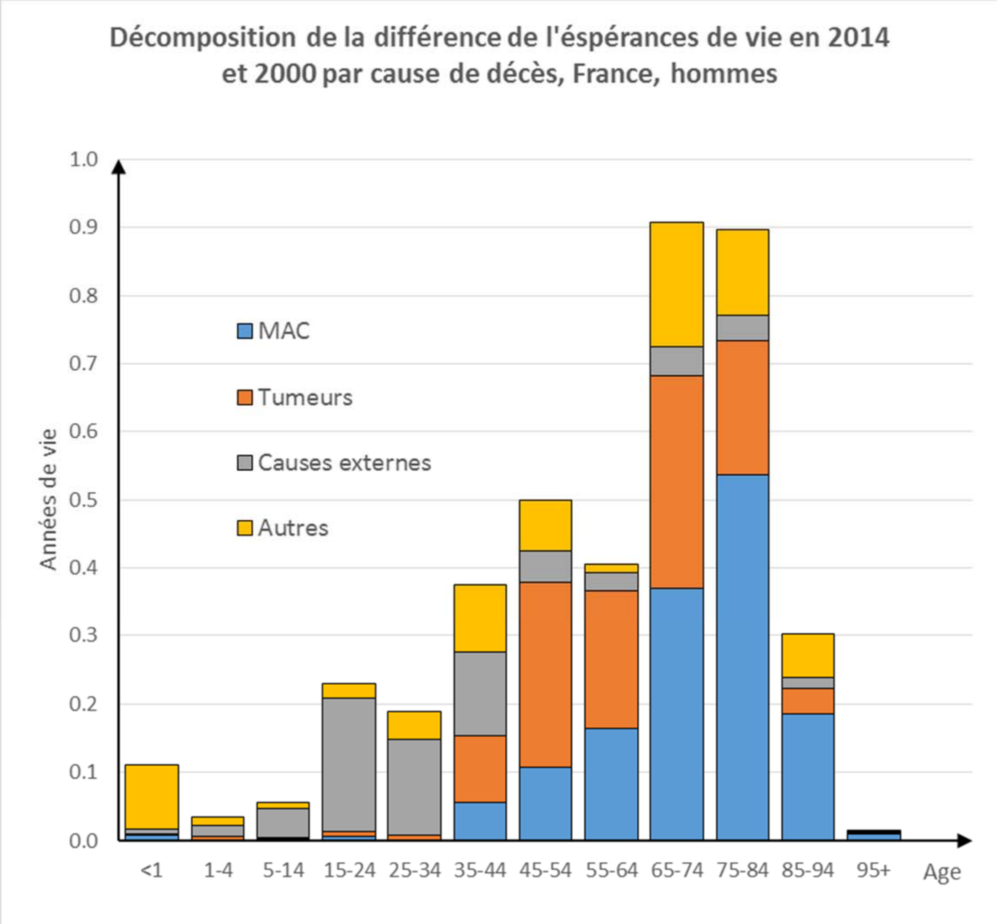
$$ {}_n\Delta_x^i = {}_n\Delta_x\cdot\frac{{}_nm_x^i(2) - {}_nm_x^i(1)}{{}_nm_x(2) - {}_nm_x(1)} \rightarrow $$

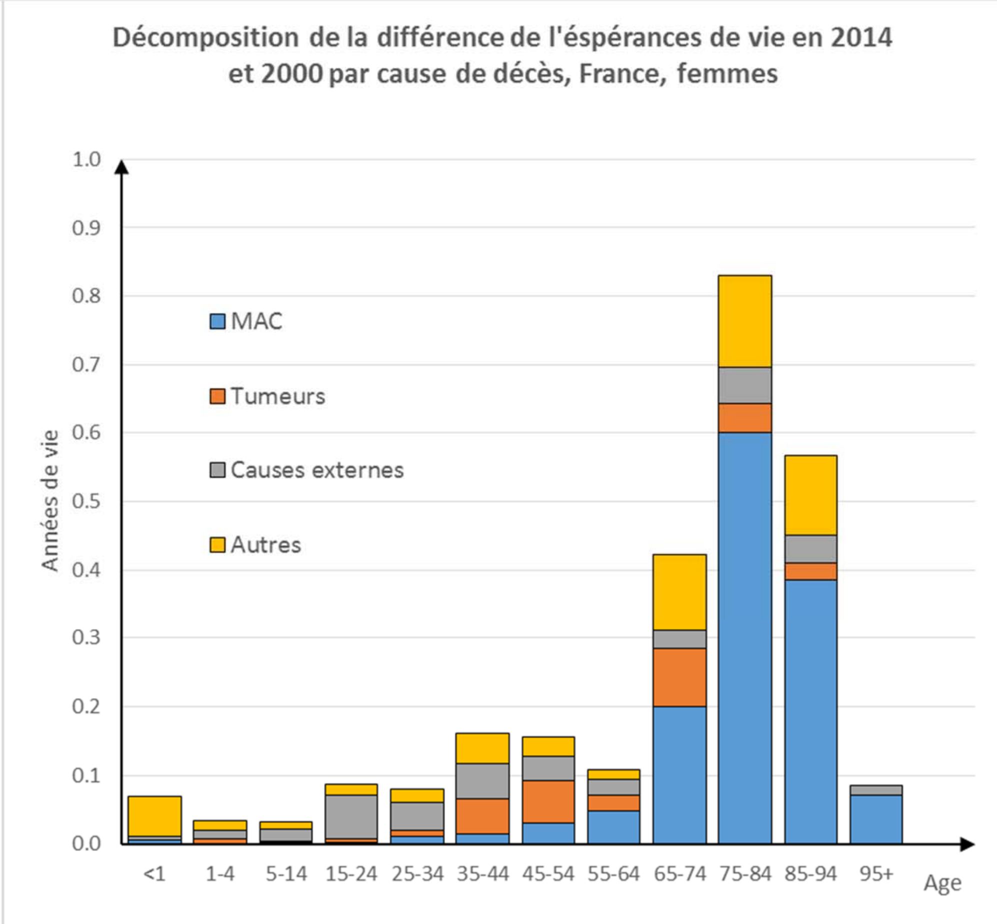
$$ {}_n\Delta_x^i = {}_n\Delta_x\cdot\frac{{}_nR_x^i(2)\cdot{}_nm_x(2) - {}_nR_x^i(1)\cdot{}_nm_x(1)}{{}_nm_x(2) - {}_nm_x(1)} $$

Les composants de la variation de l'espérance de vie par âge et par causes de décès (exemple France, 2000 vs 2014)

$e_0(2014) = 79.26$; $e_0(2000) = 75.24$

$e_0(2014) = 85.4$; $e_0(2000) = 82.8$





$e_0(2014) - e_0(2000) = 4.02 =$
1.45(MAC) + 1.14 (tumeurs) + 0.69 (causes externes) +
+ 0.74 (autres)
ou 36% + 28% + 17% +18%

$e_0(2014) - e_0(2000) = 2.60 =$
1.37(MAC) + 0.31 (tumeurs) + 0.39 (causes externes) +
+ 0.53 (autres)
53% + 12% + 15% + 21%

Estimation de la variance dans une table de mortalité

La survie est un processus stochastique discret de type Bernoulli avec une variable aléatoire qui ne prend que deux valeurs (vivant vs décès) avec :

p_i – la probabilité que l'individu reste en vie sur l'intervalle entre x_i et x_{i+1}

q_i – la probabilité que l'individu décède sur l'intervalle entre x_i et x_{i+1}

les espérances mathématiques d'une survie et d'un décès qui sont les événements complémentaires et disjoints distribués selon la loi binominale : $\boldsymbol{p_i + q_i = 1}$

avec la variance sur l'échantillon pour la proportion de survivants $\boldsymbol{S^2_{p_i}}$ et celle de la proportion de décédés $\boldsymbol{S^2_{q_i}}$ qui sont égales $\boldsymbol{S^2_{p_i}} = \boldsymbol{S^2_{q_i}}$

Soit $\boldsymbol{D_i}$ – le nombre de décès sur l'intervalle *(x, x+1)* et $\boldsymbol{N_i}$ – le nombre de vivants à l'âge xi, et la probabilité de décéder : $\boldsymbol{q_i} = \frac{\boldsymbol{D_i}}{\boldsymbol{N_i}}$

si tous les individus en vie sur l'intervalle ***i*** d'un échantillon $\boldsymbol{N_i}$ ont la même probabilité de décéder durant cet intervalle, alors $\boldsymbol{D_i}$ est une variable aléatoire binomiale et $\boldsymbol{q_i}$ est la proportion binomiale avec leurs variances :

$$\boxed{S^2_{D_i} = N_i \cdot q_i \cdot (1-q_i)} \rightarrow S^2_{D_i} = D_i \cdot (1-q_i) \quad \text{et} \quad S^2_{q_i} = \frac{q_i \cdot (1-q_i)}{N_i} \rightarrow \boxed{S^2_{q_i} = \frac{q_i^2 \cdot (1-q_i)}{D_i}}\ ^*$$

soit $\boldsymbol{a_i}$ est la fraction du dernier intervalle vécu, et $q_i = \frac{n_i \cdot M_i}{1+(1-a_i)\cdot n_i \cdot M_i}$ (voir chap.5, diap.24)

on obtient formule $S^2_{q_i} = \frac{n_i^2 \cdot M_i \cdot (1-a_i \cdot M_i)}{P_i \cdot [1+(1-a_i)\cdot n_i \cdot M_i]^3}$ → si les intervalles d'âge sont un an (n_i=1), on remplace i par x

pour obtenir les formules correspondant aux âges : $S^2_{q_x} = \frac{M_x \cdot (1-a_x \cdot M_x)}{P_x \cdot [1+(1-a_x)\cdot M_x]^3}$ et $S^2_{q_x} = \frac{4M_x \cdot (2-M_x)}{P_x \cdot [2+M_x]^3}$ si $\boldsymbol{a_x}$ = 0,5

---

*) $\boldsymbol{D_i}$ – est directement observable, $\boldsymbol{N_i}$ - non

Variance de l'espérance de vie dans un échantillon

Soit p_{ij} – la probabilité de survie dans l'intervalle (x_i, x_j) calculée comme il suit

$p_{ij} = p_i \cdot p_{i+1} \cdot \ldots \cdot p_{j-1}$ avec *i < j; i,j = 0, 1, 2, …* et $p_i = 1 - q_i$, *i = 0, 1, 2, …*

Alors la variance de p_{ij} se fabrique avec la transformation de König-Huygens

$$S^2_{p_{ij}} = p^2_{ij} \cdot \sum_{h=i}^{j-1} p_h^{-2} \cdot S^2_{p_h}$$ sachant que $S^2_{p_i} = S^2_{q_i}$ cf. diapo précèdent

On peut ensuite calculer variance de l'espérance de vie des survivant à l'âge α

$$S^2_{e_\alpha} = \sum_{i=\alpha}^{\omega-1} p^2_{\alpha i} \cdot [e_{i+1} + (1-a_i) \cdot n_i]^2 \cdot S^2_{p_i}$$

En quatre étapes (quatre colonnes) :

1. Calculer la variance de p_x (ou q_x) ; x – l'âge dans la première colonne de la table
2. Calculer le produit $l_x^2 \cdot [e_{x+1} + (1-a_x) \cdot n_x]^2 \cdot S^2_{p_x}$
3. Additionner les produits calculés dans la colonne 2 de bas vers haut
4. Diviser les sommes dans la colonne 3 par l_x^2 pour obtenir le score de la variance

Table de mortalité abrégée, France métropolitaine, Hommes, 2013

Intervalle d'âge (ans) de x_i à x_{i+1}	Amplitude de l'intervalle	Effectif de la population exposée	Nombre de décès sur l'intervalle	Taux de mortalité	Fraction vécue dans le dernier intervalle d'âge	Quotient de mortalité sur l'intervalle
x_i x_{i+1}	n_i	P_i	D_i	M_i	a_i	q_i
0 - 1	1	389 063	1 512	0.003886	0.1	0.00387
1 - 5	4	1 591 818	293	0.000184	0.39	0.00074
5 - 10	5	2 013 549	196	0.000097	0.46	0.00049
10 - 15	5	2 018 608	197	0.000098	0.54	0.00049
15 - 20	5	1 941 524	638	0.000329	0.57	0.00164
20 - 25	5	1 907 553	1 228	0.000644	0.49	0.00321
25 - 30	5	1 897 062	1 393	0.000734	0.5	0.00366
30 - 35	5	1 976 661	1 747	0.000884	0.52	0.00441
35 - 40	5	1 987 397	2 418	0.001217	0.54	0.00607
40 - 45	5	2 203 383	4 257	0.001932	0.54	0.00962
45 - 50	5	2 167 438	6 879	0.003174	0.54	0.01575
50 - 55	5	2 086 298	10 793	0.005173	0.53	0.02556
55 - 60	5	1 973 080	16 118	0.008169	0.52	0.04006
60 - 65	5	1 925 166	22 274	0.011570	0.52	0.05629
65 - 70	5	1 558 137	24 424	0.015675	0.52	0.07553
70 - 75	5	1 070 195	24 219	0.022630	0.51	0.10721
75 - 80	5	933 079	33 742	0.036162	0.51	0.16609
80 - 85	5	710 234	45 574	0.064168	0.48	0.27496
85 - 90	5	393 858	46 408	0.117829	0.45	0.44496
90 - 95	5	143 114	29 719	0.207660	0.41	0.64387
95+		184 650	7 398	0.300000		1

1) Score de a_i est estimé à partir moyenne des tables complètes supposant que ${}_1a_x$=0.5

2) Score de q_i est estimé avec la formule

$$q_i = \frac{n_i \cdot M_i}{1 + (1 - a_i) \cdot n_i \cdot M_i}$$

Estimation de la variance sur e_x dans la table de mortalité (France, hommes, 2013)

Intervalle d'âge (ans) de x_i à x_{i+1}	Nombre de vivants à l'âge x_i	Nombre de décès sur sur l'intervalle	Nombre d'années vecues sur l'intervalle	Nombre d'années vécues après l'âge x_i		Variance de p_i = var(q_i)	Éléments des calculs		Variance estimée de e_i	Écart type estimé de e_i
							Élément 1	Élément 2		
	l_i	d_i	L_i	T_i	e_i	$10^8 Sp_i^2$	f_i	F_i	$10^4 Se_i^2$	Se_i
0-1	100 000	387	99 651	7 879 016	78.79	0.98808	616 602.45	6 183 796.20	6.1838	0.025
1-5	99 613	73	398 272	7 779 365	78.10	0.18471	107 520.85	5 567 193.75	5.6106	0.024
5-10	99 539	48	497 566	7 381 093	74.15	0.12073	61 819.89	5 459 672.90	5.5103	0.023
10-15	99 491	49	497 343	6 883 526	69.19	0.12075	52 889.06	5 397 853.01	5.4532	0.023
15-20	99 442	163	496 861	6 386 183	64.22	0.42184	157 626.65	5 344 963.95	5.4051	0.023
20-25	99 279	319	495 582	5 889 322	59.32	0.83823	268 938.86	5 187 337.30	5.2629	0.023
25-30	98 960	363	493 894	5 393 739	54.50	0.96060	256 286.84	4 918 398.44	5.0223	0.022
30-35	98 597	435	491 944	4 899 845	49.70	1.10818	241 066.46	4 662 111.60	4.7957	0.022
35-40	98 163	595	489 444	4 407 901	44.90	1.51272	262 810.91	4 421 045.14	4.5881	0.021
40-45	97 567	938	485 678	3 918 458	40.16	2.15187	293 082.89	4 158 234.23	4.3682	0.021
45-50	96 629	1 522	479 643	3 432 780	35.53	3.55106	368 794.83	3 865 151.34	4.1395	0.020
50-55	95 107	2 431	469 821	2 953 137	31.05	5.89645	453 063.66	3 496 356.51	3.8654	0.020
55-60	92 676	3 713	454 470	2 483 316	26.80	9.55744	521 510.29	3 043 292.85	3.5433	0.019
60-65	88 964	5 007	432 800	2 028 845	22.81	13.42307	487 000.00	2 521 782.56	3.1863	0.018
65-70	83 956	6 342	404 561	1 596 046	19.01	21.59531	479 650.34	2 034 782.56	2.8868	0.017
70-75	77 615	8 321	367 686	1 191 485	15.35	42.36914	524 740.42	1 555 132.22	2.5816	0.016
75-80	69 294	11 509	318 270	823 799	11.89	68.17994	410 549.79	1 030 391.79	2.1459	0.015
80-85	57 784	15 889	247 611	505 528	8.75	120.28025	307 923.67	619 842.00	1.8564	0.014
85-90	41 896	18 642	158 213	257 917	6.16	236.79822	205 863.26	311 918.33	1.7771	0.013
90-95	23 254	14 972	72 100	99 704	4.29	496.78724	106 055.07	106 055.07	1.9613	0.014
95+	8 281	8 281	27 605	27 605	3.33					

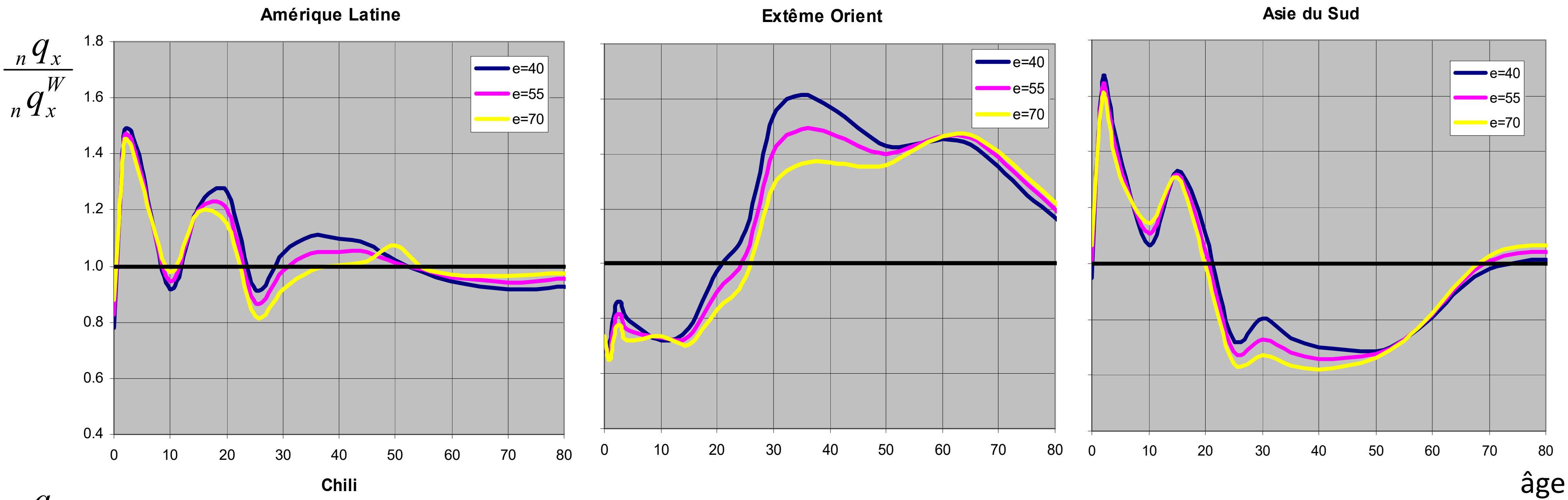
$$f_i = l_i^2 \cdot [e_{i+1} + (1 - a_i) \cdot n_i]^2 \cdot S_{p_i}^2$$

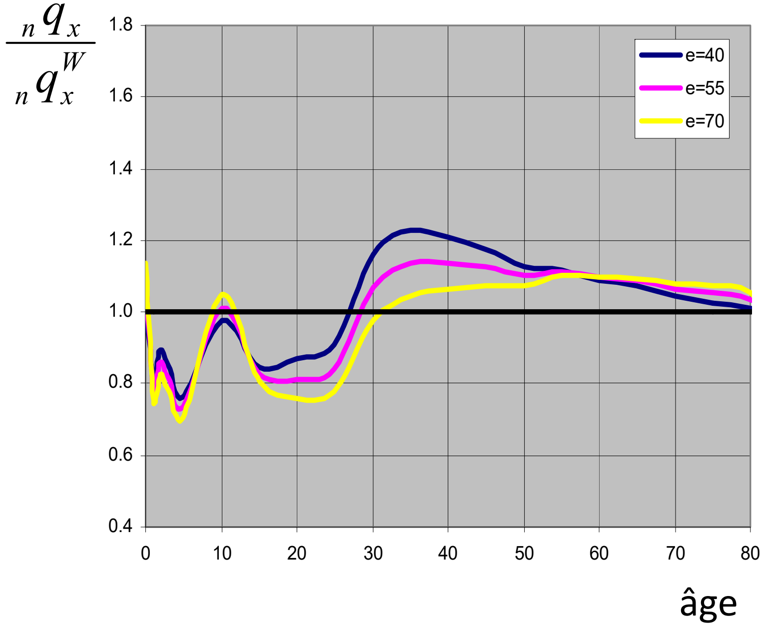
$$S_{\hat{q}_i}^2 = \frac{q_i^2 \cdot (1 - q_i)}{D_i}$$

$$F_i = \sum_{j=i}^{\omega} f_j$$

Annexe 1 :

Comparaison des tables régionales de mortalité : Tables des Nations Unies 1982 vs tables « Ouest » de C.-D.1983 (niveau 9, 15, 21)





Les autres statistiques à appliquer pour la description des tables de mortalité (du régime de mortalité), notamment les statistiques de la forme de distribution et de la concentration :

- Mode et médiane;
- Quartiles, variation interquartile;
- Indice de Gini
- Etc…

De nos jours, c'est une question qui reste assez peu étudiée (il n'y a pas de solutions standards)

Annexe 2 :

La démonstration pour la diapositive 28 (Tables multiple)

En retenant l'hypothèse que la durée de vie dans une table associée à un facteur unique (TAFU) est une moyenne entre l'amplitude de l'intervalle n (si la mortalité de cause i est absente ${}_nm_x^i = 0$) et ${}_ne_x^*$ – la durée de vie dans une TAFU (si la cause i est la seule cause de décès absente ${}_nm_x^{-i} = 0$) pondérées pas la proportion de décès de la cause i :

$$ {}_ne_x^* = \frac{{}_n^*L_x^i}{{}^*l_x^i} = n \cdot (1 - R^i) + R^i \cdot \frac{{}_nL_x}{l_x} $$ F 1.1 (4)†

L'équation 1.1 peut être simplifiée : $\frac{{}_n^*L_x^i}{{}^*l_x^i} = n - R^i \left(n - \frac{{}_nL_x}{l_x} \right)$ F 1.2

Sachant que dans toute table de mortalité ${}_ne_x$ est toujours une moyenne entre n (la durée de vie des survivants) et ${}_na_x$ – la durée moyenne de vie dans le dernier intervalle d'âge (la durée de vie des décédés entre x et $x+n$), respectivement pondérées par les probabilités de survie ${}_np_x$ et de décès ${}_nq_x$, on peut en déduire :

$$ \frac{{}_nL_x}{l_x} = n \cdot {}_np_x + {}_na_x \cdot {}_nq_x = n - (n - {}_na_x) \cdot {}_nq_x $$ F 1.3

En utilisant la formule 1.3 (en sa version finale en bleu) on peut transformer l'expression 1.2 comme le suit :

$$ \frac{{}_n^*L_x^i}{{}^*l_x^i} = n - R^i \left(n - \frac{{}_nL_x}{l_x} \right) = n - (n - {}_na_x^*) \cdot {}_nq_x^* = n - R^i \left(n - n - (n - {}_na_x) \cdot {}_nq_x \right) $$

Les n en rouge s'annulent et les signes « + » et « – » s'additionnent $({}_na_x^* - n) \cdot {}_nq_x^* = R^i \cdot {}_nq_x \left({}_na_x - n \right)$

et on obtient une formule définitive pour $\boxed{{}_na_x^* = n + R^i \cdot \frac{{}_nq_x}{{}_nq_x^*} \left({}_na_x - n \right)}$ F 1.4 (5)†

On se rappelle que dans le dernier intervalle : ${}_\infty^*a_x^i = {}^*e_\omega^i = \frac{e_\omega^0}{R^i}$

† entre les parenthèses les numéros des formules dans la diapositive d'origine (n°28)