REPERES - IREM. N° 114 - janvier 2019

DU DISCRET AU CONTINU : « UN CONTE DE NOEL » D'ARNAUD DESPLECHIN

Analyse de la « scène du tableau noir »

Alban REGNAULD
groupe « Didactique des probabilités », Perpignan

Ce texte est également consultable (*) *en ligne sur le portail des Irem, onglet : Repères IREM http://www.univ-irem.fr/*

Introduction

Sorti en 2008, après avoir été présenté en sélection officielle au festival de Cannes, réalisé par Arnaud Desplechin, « Un conte de Noël » étudie les relations entre les membres d'une famille bourgeoise du Nord, les Vuillard, face à la maladie de la maîtresse de maison. Les mathématiques y jouent un rôle notable. Une scène en particulier expose, de façon relativement détaillée, les calculs apparaissant de façon explicite, des tentatives de modélisation par les probabilités des conséquences du diagnostic de la maladie. Nous nous proposons d'effectuer une analyse aussi exhaustive que possible de cette scène, située entre les index 1 h 02 min 29 s et 1 h 04 min 40 s, afin d'explorer la validité scientifique des calculs et réflexions qui y sont présentés.

Les contenus des tableaux sont repris dans des encadrés ; les passages en italique décrivent les scènes aussi fidèlement que possible ; les transcriptions des répliques y figurent entre guillemets. Les termes techniques signalés par un astérisque sont explicités dans le lexique à la fin de l'article.

Le contexte de la scène.

Pour tenter de convaincre son épouse Junon de démarrer un traitement contre la myélodysplasie* qui lui a été diagnostiquée, Abel Vuillard tente de quantifier l'impact qu'aurait la décision de se soigner ou non sur son « espérance de vie ». Entrera ensuite en scène Claude Dédalus, mathématicien professionnel, lauréat de la Médaille Fields, qui reprendra et

(*) Nous remercions le producteur qui a donné son accord pour la mise en ligne de la vidéo de l'extrait du tableau noir dans Un conte de Noël de Arnaud Desplechin, afin d'accompagner la mise en ligne de l'article.

approfondira l'étude, en passant d'un modèle discret à un modèle continu.

Avant de détailler le raisonnement d'Abel, puis son approfondissement par Claude, je tiens à préciser que cette scène a été conçue avec le soutien scientifique de MM. Wendelin Werner et Cédric Villani, et je remercie chaleureusement M. Villani pour le temps qu'il m'a consacré, et les éclaircissements qu'il m'a apportés.

Le modèle discret.

Abel commence par dresser l'arbre pondéré suivant, qui résume un certain nombre d'informations distillées dans le film :

Les tests font apparaître 8,2% d'erreurs de diagnostic ou de pronostic. Le taux de survie à un an des malades non traités est de 50%. Le traitement tue 20% des personnes traitées (GVH : graft versus host disease : le greffon se retourne contre le receveur). Dans 30% des cas, le traitement échoue et on retrouve le taux de létalité de 50% par an. Dans 50% des cas, la greffe prend et le malade est guéri.

Les syndromes myélodysplasiques sont de nature et de gravité variées, pouvant évoluer ou non en leucémie, et cela rend de fait difficile la vérification de la pertinence des données médicales exposées, faute d'informations sur la nature exacte du syndrome myélodysplasique de Junon. On peut s'interroger sur ce 8,2% de faux positifs. Pourquoi ne pas multiplier les tests ? Dans la pratique, un syndrome myélodysplasique se diagnostique à l'aide plusieurs examens médicaux de différentes natures (hémogramme, ponction médullaire, etc.), nécessaires pour en déterminer la nature et la gravité, et établir le pronostic.

Abel entreprend ensuite le calcul de l'espérance de vie de son épouse dans chaque cas croisé : malade ou non, traitée ou non.

Il faut prendre ici l'expression « espérance de vie* » au sens de « espérance mathématique de durée de survie* », ce qui est différent du calcul officiel de l'espérance de vie selon les normes OMS ou INSEE. Nous parlerons donc désormais d'espérance de temps de survie. Plus loin dans la scène, le personnage de Claude

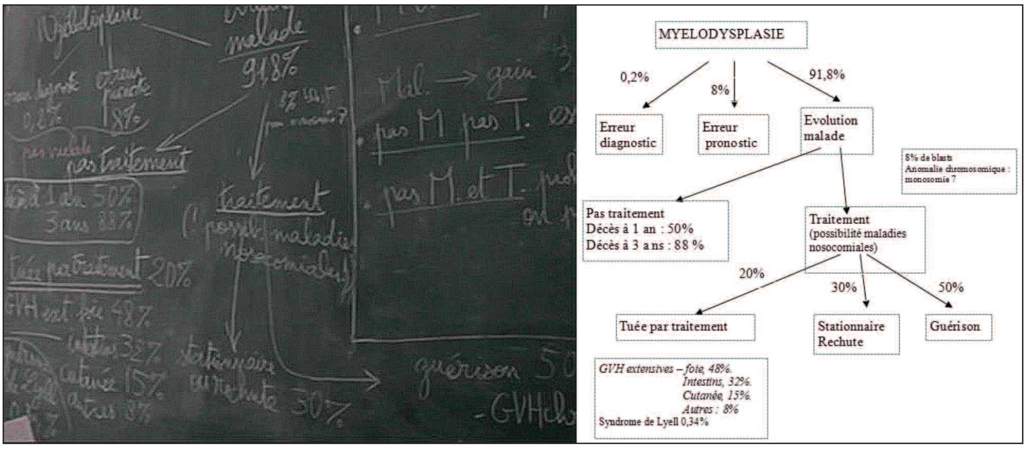
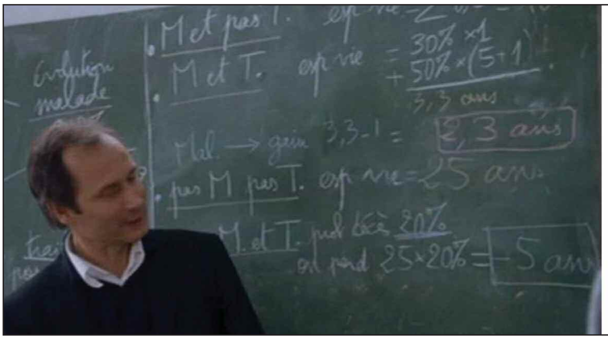


Figure 1 : l'arbre pondéré.



M et pas T esp vie = $\sum 0{,}5^n = 1$ an
M et T esp vie = 30% × 1 + 50% × (5 + 1)
= 3,3 ans
Malade gain 3,3 – 1 = 2,3 ans
pas M pas T esp vie = 25 ans
pas M et T prob décès 20%
On perd 25×20% = - 5 ans

Figure 2 : Tableau des cas croisés.

Dédalus use de cette ambigüité lorsqu'il dit *« plus on vit longtemps, plus l'espérance de vie s'accroît »*, phrase dont on imagine aisément les débats qu'elle pourrait déclencher dans une salle de classe. Bien sûr, plus un individu vit longtemps, plus l'espérance de vie de la population à laquelle il appartient s'accroît, mais plus l'espérance de temps de survie de l'individu, a priori, diminue.

Le cas « non malade et non traitée » donne une espérance de temps survie de 25 ans, basée justement sur l'espérance de vie donnée par les pouvoirs publics, compte tenu de l'âge du personnage joué par Catherine Deneuve.

Le cas « non malade et traitée » donne une espérance de temps de survie de 20 ans, en tenant compte des 20% de patients traités qui succombent à la greffe : 0,20 × 0 + 0,80 × 25 = 20*. D'où une perte de 5 années d'espérance de temps de survie, résumée par cette ligne pour le moins inadéquate* « on perd 25 × 20% = – 5 ans ».

Le cas « malade et pas traitée » est détaillé sur le « paperboard » ci-contre :

Abel étudie la suite du nombre de survivants au bout de n années. Le taux de létalité au bout d'un an étant de 50%, et celui au bout de trois

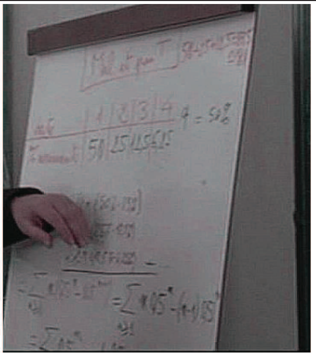


Figure 3a : le « paperboard » où est détaillé le modèle discret de la loi de temps de survie des malades non traités.

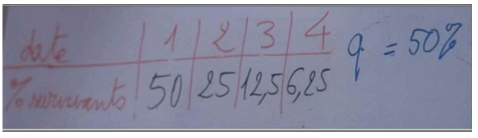


Figure 3b : détail du paperboard

Figure 3c : détail du tableau : le taux de létalité à 1 an et 3 ans.

ans de 88%, il considère une suite géométrique de raison 0,5. Remarquons que

$100\% - (50\%)^3 = 100\% - 12{,}5\% = 87{,}5\% \approx 88\%.$

Il calcule l'espérance de temps de survie en effectuant le calcul :

$$\begin{aligned}\mu &= 0 \times (50\% - 25\%) + 1 \times (25\% - 12{,}5\%) \\ &\quad + 2 \times (12{,}5\% - 6{,}25\%) + \ldots \\ &= \sum_{n \geq 0} n\,(0{,}5^n - 0{,}5^{n+1}) \\ &= \sum_{n \geq 1} (n\,0{,}5^n - (n-1)0{,}5^n) \\ &= \sum_{n \geq 1} 0{,}5^n = 1\ .\end{aligned}$$

On peut remarquer que la première ligne, décomposée avec les points de suspension, est fausse et non cohérente avec les notations sommées suivantes qui, elles, sont exactes. Le « bon » calcul eût plutôt été :

$$\begin{aligned}\mu = 0 \times (100\% - 50\%) &+ 1 \times (50\% - 25\%) \\ &+ 2 \times (25\% - 12{,}5\%) \\ &+ 3 \times (12{,}5\% - 6{,}25\%) + \cdots\end{aligned}$$

Abel en conclut que sans traitement, l'espérance de temps de survie d'une personne malade est de 1 an.

Pour le cas « malade et traitée », il effectue la moyenne pondérée des espérances de survie dans les situations décrites sur l'arbre pondéré : tuée par le traitement (espérance de temps de survie : 0 *; probabilité* 20%*), rechute = traitement inefficace, donc équivalent à* **« malade et non traitée »** *(espérance de survie :* 1 an *d'après le calcul précédent ; probabilité 30%), et guérison : (probabilité 50% ; espérance de survie* 5 + 1*. Abel considère qu'à l'année d'espérance de temps de survie des malades non traités s'ajoutent les cinq années de vie « récupérées » au cas « non malade et traité », dont on a vu plus haut qu'il faisait perdre cinq années d'espérance de temps de survie. Ces cinq années de vie sont donc « non perdues » du fait de la guérison et du non-rejet de la greffe.) D'où :*

$$0{,}2 \times 0 + 0{,}3 \times 1 + 0{,}5 \times (5+1) = 3{,}3 \text{ ans.}$$

En retirant l'année d'espérance de temps de survie d'un malade non traité, il en conclut que le traitement fait gagner 2,3 *années d'espérance de temps de survie si Junon est malade.*

Cette partie du calcul effectué par Abel est sans doute contestable : il aurait pu paraître tout aussi pertinent de considérer que si le non traitement ou son absence de résultat donne une espérance de survie de 1 an, la guérison totale entraîne la « récupération » des 25 années d'espérance de survie posées en préalable par Abel.

$$0{,}2 \times 0 + 0{,}3 \times 1 + 0{,}5 \times 25 = 12{,}8\ .$$

Le résultat serait alors très différent, ce qui doit nous rappeler que les choix de modélisation ne sont pas sans conséquence sur les éventuelles conclusions qu'on peut tirer d'une étude.

Le modèle continu.

Entre alors en scène Claude Dédalus, lauréat de la médaille Fields et gendre des Vuillard.

En étudiant le travail d'Abel, il fait remarquer que la modélisation par une suite géométrique du nombre de survivants fait implicitement l'hypothèse que « les décès se produisent à la date anniversaire » du début de l'étude.

En effet, la modélisation d'Abel revient à considérer que les 50% qui survivent moins d'un an meurent tous au jour du début de la maladie, que ceux qui survivent plus d'un an mais moins de deux meurent tous exactement un an après le début de la maladie, etc.

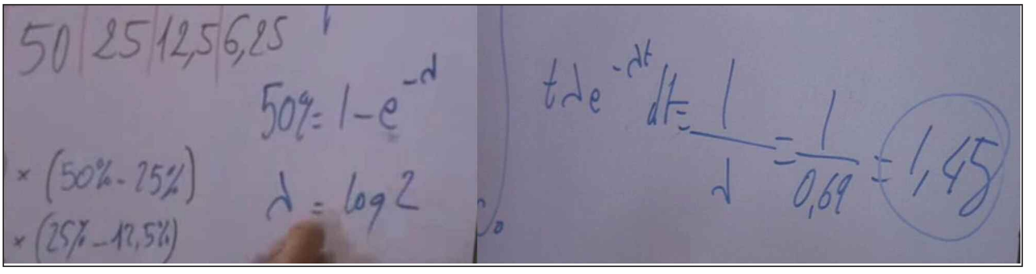


Figure 4 : La modélisation de la loi de probabilité du temps de survie par une loi exponentielle.

Claude propose alors de remplacer ce modèle discret pour le moins approximatif par un modèle continu : une loi exponentielle. Partant de l'hypothèse $p(X \leq 1) = 0{,}5$ *, il en déduit le paramètre de cette loi :* $\ln 2$.

On remarquera la notation « log 2 » qui, du point de vue des élèves de Terminale d'aujourd'hui, permet aisément de confondre ce logarithme népérien avec un logarithme décimal. Cependant, les mathématiciens professionnels emploient couramment *log* pour *ln*, et log_{10} pour le logarithme décimal.

Claude calcule ensuite l'espérance de cette loi exponentielle qui vaut $\dfrac{1}{ln(2)} \approx 1{,}442$,

arrondi à 1,45 *...*

Le passage du modèle discret au modèle continu fait donc gagner *« presque six mois de vie »*, ce qui s'explique aisément du point de vue des calculs de surface : la surface des rectangles étant remplacée par celle sous la courbe.

En se remémorant que le but avoué des protagonistes de la scène est de convaincre Junon d'entreprendre le traitement, on pourrait remarquer que si Abel avait modélisé la suite géométrique des aires des rectangles situés au-dessus de la courbe, c'est-à-dire en considérant que les malades meurent à la fin de chaque période et non au début, il aurait obtenu une espérance plus haute que celle du modèle continu…

Le calcul de la moyenne pondérée des espérances des scénarios « malade et traitée » devient donc :

$0{,}2 \times 0 + 0{,}3 \times 1{,}45 + 0{,}5 \times (5 + 1{,}45) = 3{,}66$,

arrondi à 3,7 *ans.*

Le calcul de l'espérance de gain d'espérance de temps de survie en cas de traitement effec-

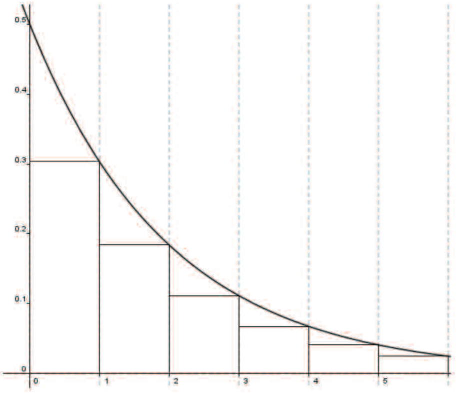


Figure 5, comparaison des modèles discret et continu.

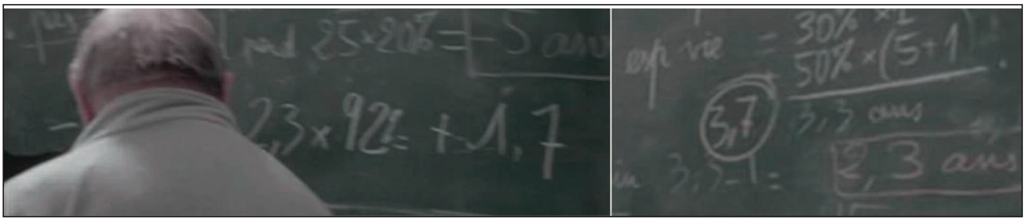


Figure 6, les derniers calculs de Claude et Abel.

tué par Abel (3,3 – 1) devenant alors (3,7 – 1,45) donnant 2,25 arrondi à 2,3 . Ce qui au final, du fait des arrondis, ne change quasiment rien ! Ce calcul n'est pas exposé dans le film.

Enfin, il calcule la moyenne pondérée des 5 *années perdues en espérance de temps de survie par les* 8% *de patients non malades, et les* 2,3 *années gagnées par les malades traités :*

$$-5 \times 0.08 + 2,3 \times 0,92 \approx 1,7 \ .$$

Il est à remarquer que ce calcul aurait pu être effectué par Abel avec les chiffres de sa modélisation discrète.

Par ailleurs, le passage du discret au continu, pour pertinent qu'il soit intrinsèquement, semble ne pas changer de manière significative le résultat d'un calcul qui eût pu être fait par Abel sans l'intervention de Dédalus. Et même, le passage du discret au continu, en donnant une espérance de temps de survie sans traitement supérieure à la modélisation discrète, pourrait nuire au propos des protagonistes, en diminuant d'autant le gain potentiel du traitement, et ce serait le cas s'ils effectuaient le calcul en valeurs exactes et non approchées !

Dédalus en conclut, pour finir, que « malade ou non, si Junon se soigne, en moyenne, elle gagne à peu près deux ans (d'espérance de temps de survie). »

Conclusion.

Cette scène présente un intérêt pédagogique certain, et ce travail d'analyse a été fait dans le but de fournir les outils permettant de proposer un scenario de séquence en fin de terminale S, ou éventuellement en postbac. Le scénario envisagé doit permettre de réinvestir, dans un contexte motivant et peu habituel, les connaissances acquises en première et terminale. Probabilités conditionnelles, problématique des tests, loi binomiale, lois continues et en particulier la loi exponentielle, la notion d'espérance mathématique, les suites et séries géométriques, l'approche de l'intégrale par la méthode des rectangles ; le nombre de notions abordées en deux minutes de film est considérable.

Ce travail doit faire prendre conscience aux élèves du potentiel de l'utilisation des probabilités comme outil d'analyse et de prise de décision dans de nombreux domaines professionnels. Pour le déroulement de la séquence, on propose, après le visionnage de la scène, une réflexion autour d'un document récapitulant les différents tableaux extraits du film, visant à permettre aux élèves de reconstituer, comme on l'a fait, les différentes étapes des raisonnements des protagonistes, en identifiant les lois mises en jeu par les deux modèles proposés et leur pertinence. Dans un second temps, un regard critique doit permettre de détecter les erreurs commises. L'usage discutable d'arrondis, les amalgames entre espérance de vie, espérance de temps de survie et espérance de gain d'espérance de temps de survie, les erreurs d'écritures dans les calculs sont autant d'occasions de réfléchir aux pièges qui guettent l'apprenti mathématicien manquant de rigueur dans ses calculs et ses définitions.

Dans le bilan, on pourra faire apparaitre que ce procédé de passage du discret au continu est courant dans de nombreux domaines, notamment en économie. Cette scène donne d'ailleurs des pistes intéressantes pour imaginer des activités pédagogiques transversales, en lien avec des disciplines aussi variées que la Géographie (espérance de vie), les Sciences de la Vie et de la Terre (problématique des tests, du diagnostic et du pronostic en médecine), les options Cinéma-Audiovisuel, etc. Par ailleurs, l'élaboration avec la classe des feuilles de calcul présentées dans le Lexique à la rubrique « espérance de vie » fournit une activité de simulation à même d'aider les élèves à mieux saisir les concepts complexes mis en jeu dans la situation du calcul institutionnel de l'espérance de vie.

Cette situation va être proposée pour alimenter la réflexion des enseignants qui doivent animer un laboratoire de mathématiques en lycée à Perpignan.

LEXIQUE

Myélodysplasie :

Les **syndromes myélodysplasiques** sont des maladies de la moelle osseuse. Les trois lignées cellulaires sanguines (globules rouges, globules blancs et plaquettes) présentent un trouble de différenciation qui aboutit à une production insuffisante d'une, deux ou des trois sortes de cellules. De plus la moelle osseuse produit des cellules anormales dites *« myélodysplasiques »*. Il s'agit d'une affection du sujet âgé, dont la médiane d'âge au moment du diagnostic se situe aux environs de 70 ans. Chez l'adulte, 8 à 10 % des cas surviennent en dessous de 50 ans. Les patients atteints de syndromes myélodysplasiques se répartissent, selon leur probabilité de développer une leucémie, en groupe à faible risque, à risque intermédiaire 1, à risque intermédiaire 2, ou à haut risque. Aucun indice ne permet dans le film de savoir dans quelle catégorie entre Junon.

Espérance de vie :

L'espérance de vie à la naissance (ou à l'âge 0*) représente la durée de vie moyenne — autrement dit l'âge moyen au décès — d'une génération fictive soumise aux conditions de mortalité de l'année. Elle caractérise la mortalité indépendamment de la structure par âge. (définition de l'*INSEE*)*

Cette statistique est calculée sous l'égide de l'ONU, et publiée par de nombreux organismes, incluant l'OMS. Le calcul de l'espérance de vie à la naissance d'une population s'opère en deux phases :

— Dans un premier temps, on détermine la probabilité de décéder à chaque âge (la probabilité de mourir à 1 an, à 2 ans, etc.). Pour cela, on rapporte le nombre de personnes décédées à un âge donné au nombre de personnes ayant cet âge dans la population l'année considérée. On obtient ainsi le taux de mortalité par âge.

— Dans un second temps, on imagine une génération fictive de 1 000 personnes, à laquelle on retire le nombre de personnes correspondant au taux de mortalité. Par exemple, si le taux de mortalité à 1 an est de 4‰, et celui à 2 ans de 2‰, on retire d'abord 4 personnes, puis 0,2% des personnes restantes, et ainsi de suite, jusqu'à extinction de la population. On calcule alors la moyenne des âges de décès observés.

L'âge obtenu ne dépend pas de la structure de la population étudiée (jeune ou vieillissante…) et autorise donc des comparaisons entre des populations aux pyramides des âges très différentes.

Pour illustrer cette définition, on a simulé les statistiques d'une population d'une espèce dont la durée de vie n'excède pas 10 ans. Les effectifs par âge de la population ainsi que le nombre de décès sont tirés au sort. Ainsi, sur l'exemple ci-dessous (figure 6a), il y a 362 individus de moins d'un an, dont 16 sont décédés, soit un taux de mortalité à moins d'un an de 4,42%. Dans la population fictive de 1000 individus, on considèrera donc 44 décès la première année. Il y a ensuite 558 individus âgés d'un an, dont 105 décèdent avant leur deuxième anniversaire, soit un taux de mortalité de 18,82%. On retirera donc des 956 survivants de la population fictive 18,82% de 956 soit 180 individus. Il y a ensuite 572 individus âgés de 2ans, dont 105 sont morts dans l'année, soit un taux de mortalité de 6,82%. Dans la population fictive correspondante, où il reste 776 individus de 2 ans et plus, cela représente donc 53 décès à l'âge de 2 ans. Et ainsi de suite.

	POPULATION OBSERVEE					POPULATION FICTIVE			
âge n	population au début de l'année d'âge = n	décès durant l'année	taux de mortalité			âge n	population d'âge >= n	nombre de décès	
0	362	16	4,42%			0	1000	44	
1	558	105	18,82%			1	956	180	
2	572	39	6,82%			2	776	53	
3	666	21	3,15%			3	723	23	
4	418	4	0,96%			4	700	7	
5	993	52	5,24%			5	694	36	
6	493	89	18,05%			6	657	119	
7	133	0	0,00%			7	539	0	
8	325	60	18,46%			8	539	99	
9	148	71	47,97%			9	439	211	
10	65	65	100,00%			10	228	228	
						11	0	0	
effectif total	4733								
		âge moyen de décès							
		5,41				espérance de vie		6,25	année

Figure 6a : un exemple de simulation.

	POPULATION OBSERVEE				POPULATION FICTIVE			
âge	population au début de l'année	décès durant l'année	taux de mortalité		âge	population	décès	
0	4957	149	3,01%		0	1000	30	
1	4059	78	1,92%		1	970	19	
2	3429	130	3,79%		2	951	36	
3	3130	193	6,17%		3	915	56	
4	2723	44	1,62%		4	859	14	
5	2503	107	4,27%		5	845	36	
6	3015	460	15,26%		6	809	123	
7	1626	289	17,77%		7	685	122	
8	1994	353	17,70%		8	564	100	
9	780	226	28,97%		9	464	134	
10	222	222	100,00%		10	329	329	
					11	0	0	
		âge moyen de décès						
		5,99			espérance de vie		7,39	année

Figure 6b : exemple de simulation.
Population jeune, faible mortalité infantile,
forte espérance de vie à la naissance.

	POPULATION OBSERVEE				POPULATION FICTIVE			
âge n	population au début de l'année d'âge = n	décès durant l'année	taux de mortalité		âge n	population d'âge >= n	nombre de décès	
0	1686	25	1,48%		0	1000	15	
1	1346	21	1,56%		1	985	15	
2	1471	103	7,00%		2	970	68	
3	2015	100	4,96%		3	902	45	
4	2838	59	2,08%		4	857	18	
5	2982	276	9,26%		5	839	78	
6	2159	78	3,61%		6	762	28	
7	2165	243	11,22%		7	734	82	
8	1939	269	13,87%		8	652	90	
9	998	425	42,59%		9	561	239	
10	126	126	100,00%		10	322	322	
					11	0	0	
effectif total	19725							
		âge moyen de décès						
		6,70			espérance de vie		7,58	année

Figure 6c : exemple de simulation.
Population vieillissante, faible mortalité infantile,
forte espérance de vie à la naissance.

	POPULATION OBSERVEE				POPULATION FICTIVE			
âge	population au début de l'année	décès durant l'année	taux de mortalité		âge	population	décès	
0	4064	961	23,65%		0	1000	236	
1	4313	994	23,05%		1	764	176	
2	3571	493	13,81%		2	588	81	
3	3459	758	21,91%		3	506	111	
4	3149	440	13,97%		4	395	55	
5	2991	324	10,83%		5	340	37	
6	2808	1162	41,38%		6	303	126	
7	1989	815	40,98%		7	178	73	
8	1954	238	12,18%		8	105	13	
9	606	14	2,31%		9	92	2	
10	143	143	100,00%		10	90	90	
					11	0	0	
		âge moyen de décès						
		3,75			espérance de vie		3,36	année

Figure 6d : exemple de simulation.
Population jeune, forte mortalité infantile,
faible espérance de vie à la naissance.

Espérance de temps de survie :

C'est l'espérance d'une variable aléatoire mesurant le temps de vie avant décès d'un élément d'une population. La loi de probabilité de cette variable est en général modélisée à partir d'une étude statistique, comme le fait Abel pour les malades non traités. La loi de durée de vie d'un atome radioactif, les lois de durée de vie d'un système mécanique avant panne, sont des exemples de telles lois.

SOURCES BIBLIOGRAPHIQUES

Toutes ces pages Internet ont été consultées le 02/09/2016

http://www.insee.fr/fr/methodes/default.asp?page=definitions/esperance-vie.htm

https://fr.wikipedia.org/wiki/Loi_de_fiabilité

https://fr.wikipedia.org/wiki/Analyse_de_survie

www.fr.medipedia.be/myelodysplasies

https://fr.wikipedia.org/wiki/Syndrome_myélodysplasique

www.myelodysplasies.org

Groupe Francophone des Myélodysplasies, (2015). Consensus français sur les syndromes myelodysplasiques (smd) et la leucémie myélomonocytaire chronique :

http://www.gfmgroup.org/fichiers/ConsensusGFM2015.pdf

Rodríguez, G. (2007). Ch 7 : Survival Analysis, in *Lecture Notes on Generalized Linear Models*,

http://data.princeton.edu/wws509/notes/