

pas être déterminée. On aimerait quand même avoir une idée de cette espérance de vie, y compris pour les gens nés en 2021 ou en 1990! Comment tirer parti des statistiques passées?

Une première idée serait de calculer l'âge moyen de décès (AMD), durant l'année N , des personnes auxquelles on s'intéresse, par exemple les Françaises durant l'année 2021. Si la structure des âges et les taux de mortalité dans la population étudiée étaient constants dans le temps, cette méthode serait satisfaisante. Ce n'est pas le cas.

L'idée du calcul de l'Insee et des statisticiens dans tous les pays est plus subtile: ils calculent les taux de mortalité par âge à partir des données les plus récentes disponibles (celle de 2019 par exemple), puis ils imaginent une population fictive qui évolue d'année en année en subissant ce taux de mortalité pour chaque âge. Les enfants de moins d'un an de cette population fictive meurent en proportion comme l'indique le taux de mortalité pour 2019 des enfants de moins d'un an; puis les enfants ayant survécu à leur première année meurent dans leur seconde année avec le taux de mortalité pour 2019 des enfants âgés entre un et deux ans, etc. L'espérance de vie mathématique pour cette population fictive est alors facile à calculer: c'est la moyenne des âges de décès pondérés par les effectifs associés. Le résultat de ce calcul est ce que les statisticiens publient, et nous le nommerons «espérance de vie des statisticiens» (EV_s).

Résumons: l'espérance de vie réelle (EV_R) fondée sur la notion mathématique d'espérance (comme pour les dés) n'est calculable que pour les générations anciennes, aussi on la remplace par le calcul de EV_s , qui se fonde chaque année sur les chiffres des statistiques de décès de la dernière année pour laquelle on en dispose, en imaginant une génération fictive respectant ces taux de mortalité. Les encadrés 1 et 2 donnent des exemples montrant numériquement les différences entre EV_R et EV_s . Cela étant, quels risques prend-on en calculant EV_s à la place de EV_R ?

POPULATIONS IMAGINAIRES, MAIS POSSIBLES

L'espérance de vie à la naissance n'est pas l'information qui vous concerne le plus, c'est plutôt «l'espérance de vie à l'âge A », quand vous avez l'âge A .

Dans le cas de EV_R , ce sera l'espérance au sens mathématique du nombre d'années de vie restant à vivre pour une personne d'âge A , en se référant aux chiffres réels des décès de gens ayant l'âge A à la date à laquelle on s'intéresse (par exemple $N = 1880$), qui ne sont connus que progressivement quand les années passent.



ESPÉRANCE DE VIE ET ÂGE MOYEN DE DÉCÈS

L'espérance de vie réelle (EV_R), l'espérance de vie des statisticiens (EV_s) et l'âge moyen de décès (AMD) sont différents. Imaginons un élevage animal dont les effectifs par âge aux 1^{er} janvier 1900, 1901, 1902 et 1903 sont donnés par le tableau :

Âge (ans)	0	1	2	3	4
Effectif 1900	100	100	100	100	0
Effectif 1901	200	80	60	50	0
Effectif 1902	300	170	56	30	0
Effectif 1903	400	270	136	28	0

On en déduit les taux de décès par âge (TCA) chaque année :

Âge (ans)	0	1	2	3
TCA en 1900	20%	40%	50%	100%
TCA en 1901	15%	30%	50%	100%
TCA en 1902	10%	20%	50%	100%

Nous avons supposé que le nombre de naissances augmentait globalement, ce qui se produit en général pour les populations humaines. Ici, pour un élevage, ce nombre passe de 100 à 400. Nous avons aussi supposé que les taux de mortalité étaient chaque année croissants en fonction de l'âge, mais que, d'une année à la suivante, ils décroissaient avec les progrès dans les soins apportés aux animaux de l'élevage. Pour les humains, il y a aussi une augmentation des taux de mortalité avec l'âge et une baisse de ces taux d'année en année avec les progrès de la médecine et de l'hygiène.

Espérance de vie EV_R à la naissance pour les animaux d'âge 0 en 1900

Parmi les 100 individus nés en 1900, 20 sont morts avant d'atteindre l'âge de 1 an, 24 sont morts dans leur seconde année, c'est-à-dire quand ils avaient 1 an, 28 sont morts quand ils avaient 2 ans, et 28 quand ils avaient 3 ans. L'espérance de vie pour cette cohorte née en 1900 est donc : $EV_R = 0 \times 20/100 + 1 \times 24/100 + 2 \times 28/100 + 3 \times 28/100 = 1,64$ an. Pour la connaître, il faut attendre la

mort de tous les animaux concernés, mais c'est la véritable espérance de vie : la moyenne des âges de décès pondérés par les probabilités de chacun de ces âges de décès.

Espérance de vie EV_s à la naissance pour les animaux d'âge 0 en 1900 évaluée à partir des taux de mortalité de 1900

Si nous calculions l'espérance de vie des animaux de l'élevage à la fin de l'année 1900, nous devrions utiliser les taux de mortalité pour l'année 1900. C'est la situation des statisticiens qui, en 2021, veulent calculer l'espérance de vie des enfants qui viennent de naître en ne disposant que des données de 2020 ou même de 2019. Pour l'élevage, le statisticien fait l'hypothèse que les taux de mortalité de 1900 seront valides tout au long de la vie des animaux nés en 1900.

Les taux de mortalité en fonction de l'âge en 1900 sont 20 %, 40 %, 50 %, 100 %. La population fictive imaginée par le statisticien pour le calcul de l'espérance de vie passera donc de 100 individus à 80 entre 1900 et 1901, puis de 80 à 48 entre 1901 et 1902, puis de 48 à 24 entre 1902 et 1903, puis disparaîtra. Cela donne : $EV_s = 0 \times 20/100 + 1 \times 32/100 + 2 \times 24/100 + 3 \times 24/100 = 1,52$ an. L'espérance de vie du statisticien (EV_s) est inférieure à la réalité (EV_R), car le statisticien, avec les données de 1900, a surestimé les taux de mortalité, qui, grâce au progrès de l'élevage, ont diminué d'année en année.

Âge moyen de décès en 1900

Il y a eu 210 morts en 1900, et l'âge moyen des animaux décédés (AMD) est : $AMD = 0 \times 20/210 + 1 \times 40/210 + 2 \times 50/210 + 3 \times 100/210 = 2,09$ ans. Cet âge est encore plus élevé que EV_R , car les données pour 1900 contiennent trop d'animaux âgés, ce qui se corrige pour 1901, 1902 et 1903. Un plus grand nombre de sujets âgés augmente l'âge moyen de décès, comme un plus grand nombre de jeunes le ferait décroître : l'âge moyen de décès ne donne pas l'espérance de vie...

Dans le cas de EV_s , ce sera en se référant à la population fictive construite à partir des taux de mortalité les plus récents disponibles, par exemple ceux de l'année $N-1$.

Dans certains pays pauvres, un grand nombre de décès se produisent lors des premières années de vie. Ces morts prématurées influent beaucoup sur l'espérance de vie à la naissance, qui sera assez faible dans un tel cas. Il est assez clair que si, pour une population donnée, les taux de mortalité des premières années sont mauvais et qu'au-delà ils sont convenables, l'effet des mauvaises années sur l'espérance de vie à l'âge A au-delà des mauvaises années sera nul. C'est ce que l'on observe sur les courbes EV_b et EV_c de l'encadré 3.

Les cinq courbes du graphique B de l'encadré 3 indiquent l'espérance de vie à l'âge A pour cinq populations différentes déduites des données de l'Insee pour 2019 en jouant sur certaines hypothèses.

La première population envisagée est celle correspondant à la population fictive que les statisticiens imaginent à partir des données de mortalité des Français en 2019 calculées par l'Insee. C'est donc l'espérance de vie des statisticiens EV_s . Cela donne la courbe EV_a d'espérance de vie à l'âge A , en fonction de A . Les données indiquent que EV_a à la naissance est 82,3 ans. À 20 ans, l'espérance de vie est 62,8 (ce qui fait un âge espéré de décès de 82,8 ans).

La seconde population envisagée est une population fictive qui protégerait mal la vie de

ses jeunes. Elle est obtenue en multipliant par 20, pour les trente premières années, les taux de mortalité de la première population. Nous la notons EV_b . Le but des calculs est de voir l'effet sur l'espérance de vie de ces taux de mortalité importants chez les jeunes.

L'espérance de vie à la naissance est passée de 82,3 à 69 ans. Celle à 20 ans est passée de 62,8 à 58. Celle à 30 ans est en revanche, comme précédemment, de 53,1 ans. C'est normal, puisque au-delà de 30 ans, nous avons utilisé les mêmes taux de mortalité que pour la courbe EV_a ; or l'espérance de vie à l'âge A ne dépend que des taux de mortalité au-delà de l'âge A . Cela explique la coïncidence de EV_a avec EV_b à partir de 30 ans.

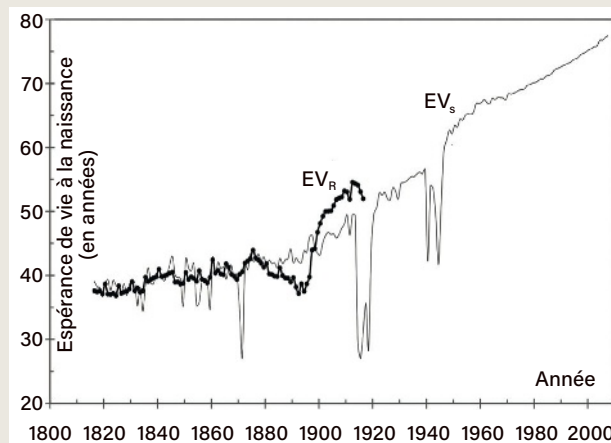
La troisième population explore encore le problème de la mortalité infantile. Cette fois, on remplace les taux de mortalité par $1/20$ pour les vingt premières années. La courbe obtenue est EV_c . L'espérance de vie à la naissance est maintenant passée à 34,7 ans, mais bien sûr à partir de 20 ans les courbes EV_a et EV_c coïncident.

Il semble logique que le nombre d'années qu'il vous reste à vivre diminue avec l'âge et c'est ce qu'on observe presque tout le temps: les courbes sont le plus souvent décroissantes. Cependant, même si cela semble paradoxal, quand le taux de mortalité est très élevé à certains âges, il se produit un renversement: l'espérance de vie augmente avec l'âge. C'est ce que l'on observe sur les courbes EV_b et EV_c . C'est

2

ESPÉRANCES DE VIE RÉELLE ET ANTICIPÉE

La courbe en trait épais représente l'espérance de vie réelle (EV_R) à la naissance pour les hommes en France, en faisant le calcul pour les générations nées entre 1816 et 1916 (et pour lesquelles le calcul est possible). La courbe en trait fin est l'espérance de vie à la naissance calculée par les statisticiens (EV_s) en n'utilisant que les données des taux de mortalité de l'année pour chacune des années de 1816 à 2007. On observe que les courbes, quand on peut calculer à la fois EV_s et EV_R , sont assez différentes. Le creux profond de EV_s vers 1915 correspond à la Première Guerre mondiale: les taux de mortalité élevés pour les classes d'âge participant aux combats créent pour le statisticien une population fictive (qui lui sert à calculer EV_s) largement défavorisée,



comparée à la réalité de l'espérance de vie (EV_R) des enfants nés aux mêmes dates. Vers 1890, la courbe EV_R est sous la courbe EV_s , et c'est aussi un effet de la guerre: en effet, les générations autour

de 1890 ont été largement victimes de la guerre (ce que EV_R prend en compte), alors que EV_s , qui n'est fondée que sur les taux de mortalité de l'année même, donc vers 1890, n'en tient pas compte.