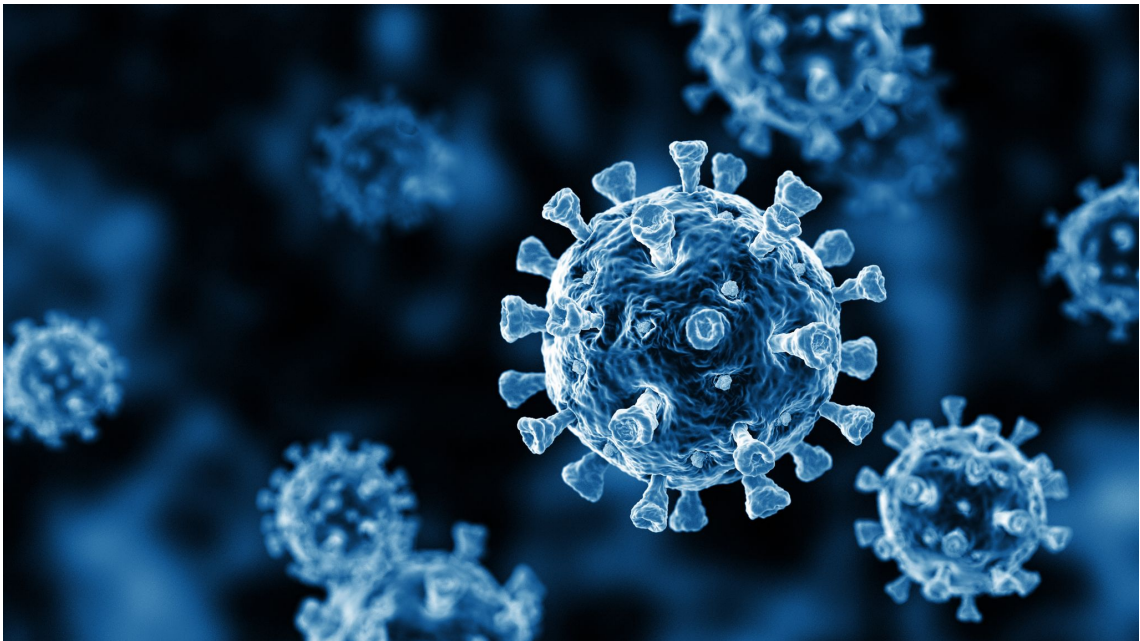


Combien de personnes la Covid-19 a-t-elle tuées en France en 2020 ?

Côme Pollet - Jules Chapon - Victor Parent - Diego Renaud

17 Mai 2023



Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 2 |
| 2 | Etude de l'article | 2 |
| 2.1 | Résumé de l'article | 2 |
| 2.2 | Incohérences et questions potentielles | 5 |
| 2.3 | A retenir | 6 |
| 3 | Modèle statistique | 7 |
| 3.1 | Définition des variables | 7 |
| 3.1.1 | Effectifs de population | 7 |
| 3.1.2 | Variables aléatoires | 7 |
| 3.1.3 | Quotients de mortalité observés | 8 |
| 3.1.4 | Quotients de mortalité prédits | 9 |
| 3.2 | Quotients de mortalité observés : théorèmes limites | 10 |
| 3.3 | Quotients de mortalité prédits : théorèmes limites | 12 |
| 3.4 | Intervalles de confiance | 13 |
| 3.4.1 | Quotients de mortalité observés | 14 |
| 3.4.2 | Quotients de mortalité prédits | 15 |
| 3.5 | Tests de surmortalité | 16 |
| 3.5.1 | Test de surmortalité à âge et sexe donnés | 16 |
| 3.5.2 | Test de surmortalité global | 17 |
| 4 | Application du modèle : Résultats | 17 |
| 4.1 | Mise en pratique du modèle | 18 |
| 4.2 | Résultats et interprétation | 19 |
| 5 | Limites du modèle et de sa mise en pratique | 20 |
| 5.1 | Limites du modèle statistique | 20 |
| 5.2 | Limites de l'application | 21 |
| 6 | Annexes | 21 |
| 6.1 | Nos Résultats | 21 |
| 6.2 | Définitions | 22 |
| 6.3 | Chiffres clés | 23 |
| 6.4 | Revue de littérature | 24 |

1 Introduction

Ce document présente nos résultats et nos analyses afin de répondre aux exigences de notre projet StatApp de 2ème année de l'ENSAE Paris, à savoir : « **Combien de personnes la Covid-19 a-t-elle tué en France en 2020 ?** ».

L'objectif de ce projet StatApp est de reproduire et critiquer l'article « L'épidémie de Covid-19 a eu un impact relativement faible sur la mortalité en France » publié par Laurent Toubiana, Laurent Mucchielli, Pierre Chaillot et Jacques Bouaud. Dans cet article, les quatre auteurs ont conduit une étude statistique aboutissant à des résultats allant à l'encontre de ceux obtenus par l'INSEE. En effet, d'après les auteurs, l'épidémie de la Covid-19 n'a pas été à l'origine d'une surmortalité chez les 0-65 ans, et n'a engendré qu'une surmortalité faible chez les plus de 65 ans. Ils obtiennent et affichent une surmortalité bien moins importante (de l'ordre de 4%) que celle obtenue par l'INSEE (près de 8%). Il semble raisonnable de vouloir reproduire leur analyse afin de vérifier la véracité de ces résultats qui vont à l'encontre de ceux présentés par la communauté scientifique, le gouvernement et les instituts d'études statistiques indépendants (INSEE, Inserm...). Nous allons tenter d'expliquer comment ces 4 auteurs ont pu aboutir à de tels résultats, et dans le même temps, essayer de quantifier l'impact de la Covid-19 sur la mortalité en France en 2020.

Pour mener à bien ce travail de Fact-Checking, nous commencerons par étudier en détails l'article en question, afin de déceler des potentielles failles dans l'analyse des auteurs. Suite à cela, nous comparerons leur méthode à celles qui font loi habituellement dans le domaine de la démographie et de l'étude de la surmortalité. Enfin, nous présenterons notre propre modèle statistique afin d'obtenir, par nous-mêmes, des estimations de la surmortalité en France en 2020.

2 Etude de l'article

L'article publié par Laurent Toubiana, Laurent Mucchielli, Pierre Chaillot et Jacques Bouaud est disponible ici.

2.1 Résumé de l'article

Dans cet article, les auteurs examinent l'impact de la pandémie de Covid-19 sur la mortalité en France en 2020, en insinuant que les mesures de confinement ont peut-être été plus délétères que salutaires. Ils commencent par un rappel historique, notant que le virus SARS-CoV-2 a émergé fin 2019 à Wuhan, en Chine, et que l'OMS a déclaré l'état de pandémie le 11 mars 2020. Ils notent que l'arrivée de l'épidémie en France vers le 15 février 2020 a surpris les autorités sanitaires en raison du nombre élevé d'hospitalisations

et de décès. Face à la crise, un confinement généralisé a été mis en place le 17 mars 2020.

En épidémiologie, la surmortalité (ou excès de mortalité) est utilisée pour mesurer l'impact d'un événement sanitaire. Les auteurs définissent la surmortalité comme **la différence entre le nombre de décès observés pendant un événement sanitaire et le nombre de décès attendus pour la même période**. Ils utilisent les données du Répertoire National d'Identification des Personnes Physiques (RNIPP), géré par l'Institut national de la statistique et des études économiques (Insee), pour obtenir des informations précises sur les décès. Ils constatent que la cause de décès n'est pas toujours clairement indiquée, ce qui peut entraîner une surévaluation de la mortalité liée à la Covid-19.

Ils soulignent **l'importance de la structure d'âge de la population** dans l'analyse de la mortalité. En raison du vieillissement de la population française, le nombre de décès augmente régulièrement. Ils concluent que l'impact de la pandémie sur la mortalité globale en 2020 doit être analysé en tenant compte de cette tendance croissante. Les auteurs de l'article ont examiné l'impact des événements sanitaires, tels que les épidémies, sur la mortalité d'une population vieillissante. Leur analyse repose principalement sur les dynamiques (plus ou moins évidentes) suivantes:

1. Les avancées médicales et les améliorations des conditions de vie ont permis un allongement de l'espérance de vie, mais cette tendance commence à ralentir, notamment en Occident.
2. Les personnes âgées sont particulièrement vulnérables aux maladies infectieuses, comme la grippe, qui peuvent causer des pics de mortalité, en particulier pendant les mois d'hiver.
3. L'article met en évidence l'existence de variations saisonnières de la mortalité, avec des pics hivernaux liés aux maladies infectieuses et des pics estivaux liés aux canicules. Par exemple, le pic mensuel le plus élevé sur une période de quinze ans a été enregistré lors de l'épidémie de grippe de l'hiver 2016-2017 avec 68 969 décès, comparativement à 67 537 en avril 2020 pendant l'épidémie de Covid-19.
4. L'effet de "moisson" ou "harvesting" est également discuté. Il s'agit d'un phénomène où les personnes âgées et fragiles meurent de manière "synchronisée" lors d'un événement sanitaire fort (épidémie, canicule), ce qui entraîne un "épuisement" temporaire de cette catégorie de population. Par exemple, la forte mortalité de 2017 a entraîné une faible mortalité en 2018 et 2019.
5. En 2020, 90% des personnes admises en réanimation et 65% des personnes décédées en lien avec l'épidémie de Covid-19 présentaient au moins une comorbidité. L'âge

médian au décès à l'hôpital des malades de la Covid-19 était de 85 ans et près de 92,5% des personnes avaient 65 ans et plus.

6. L'évolution démographique de la population française est également analysée. Les personnes âgées de 65 ans et plus, bien qu'elles ne représentent que 19,45% de la population en 2020, ont contribué à 84,3% des décès.
7. Le taux de mortalité annuel moyen pour toute la population française en 2020 était de 1,004%. Pour les individus âgés de moins de 65 ans, il était de 0,17%, mais il atteignait 4,18% pour les plus de 65 ans. Le taux de mortalité augmente de manière exponentielle en fonction de l'âge.

Pour estimer la surmortalité en 2020, les auteurs affirment avoir utilisé une méthode démographique retenue par l'**Institut National de la Santé et de la Recherche Médicale (Inserm)** et détaillée dans ce dossier. Celle-ci consiste à prendre la moyenne empirique des quotients de mortalité des 3 années précédentes (2017, 2018 et 2019). Deux estimations sont réalisées, basées sur les taux de mortalité pour les années de référence. La première se fonde sur des classes d'âge de 5 ans, tandis que la seconde, regroupe tous les âges en deux catégories : les moins de 65 ans et les 65 ans et plus. Pour chaque période, des intervalles de confiance sont calculés pour tenir compte de la variabilité des données. Toute valeur observée dans l'intervalle de confiance de la valeur prédite est considérée comme attendue, indépendamment des différences constatées.

Tableau 2 : Evaluation de la surmortalité en 2020 (3 années de référence : 2017-19) selon des classes d'âges de 5 ans

| Classe d'âge | Population en million d'habitants | Décès attendus | Décès observés | Décès en excès sur classe d'âge | Décès en excès par âge |
|--------------|-----------------------------------|-----------------------------|----------------|-------------------------------------|-------------------------------------|
| 0-4 ans | 3,52 (5,41 %) | 2 940 [2 925 ; 2 955] | 2 691 | - 234 (-7,97 % [-7,93 % ; -8,01 %]) | - 206 (-6,99 % [-6,96 % ; -7,03 %]) |
| 5-9 ans | 3,94 (6,05 %) | 289 [275 ; 303] | 272 | - 3 (-0,97 % [-0,92 % ; -1,02 %]) | - 25 (-8,79 % [-8,38 % ; -9,24 %]) |
| 10-14 ans | 4,04 (6,21 %) | 311 [293 ; 329] | 321 | 0 % | 8 (2,49 % [2,35 % ; 2,65 %]) |
| 15-19 ans | 4 (6,15 %) | 853 [812 ; 895] | 813 | 0 % | - 25 (-2,92 % [-2,79 % ; -3,07 %]) |
| 20-24 ans | 3,65 (5,6 %) | 1 425 [1 376 ; 1 474] | 1 375 | - 1 (-0,07 % [-0,06 % ; -0,07 %]) | - 30 (-2,14 % [-2,07 % ; -2,21 %]) |
| 25-29 ans | 3,63 (5,57 %) | 1 622 [1 572 ; 1 672] | 1 565 | - 7 (-0,44 % [-0,42 % ; -0,45 %]) | - 2 (-0,12 % [-0,11 % ; -0,12 %]) |
| 30-34 ans | 3,95 (6,07 %) | 2 259 [2 233 ; 2 286] | 2 255 | 0 % | - 13 (-0,59 % [-0,59 % ; -0,6 %]) |
| 35-39 ans | 4,12 (6,32 %) | 3 421 [3 383 ; 3 460] | 3 276 | - 107 (-3,12 % [-3,08 % ; -3,16 %]) | - 134 (-3,91 % [-3,87 % ; -3,96 %]) |
| 40-44 ans | 3,95 (6,07 %) | 4 758 [4 653 ; 4 864] | 4 852 | 0 % | 89 (1,87 % [1,83 % ; 1,92 %]) |
| 45-49 ans | 4,38 (6,73 %) | 8 925 [8 674 ; 9 177] | 8 679 | 0 % | - 168 (-1,88 % [-1,83 % ; -1,94 %]) |
| 50-54 ans | 4,29 (6,58 %) | 14 002 [13 706 ; 14 298] | 13 958 | 0 % | 2 (0,02 % [0,02 % ; 0,02 %]) |
| 55-59 ans | 4,23 (6,49 %) | 21 585 [20 857 ; 22 313] | 21 305 | 0 % | - 131 (-0,61 % [-0,59 % ; -0,63 %]) |
| 60-64 ans | 3,98 (6,12 %) | 30 819 [30 185 ; 31 454] | 31 116 | 0 % | - 69 (-0,22 % [-0,22 % ; -0,23 %]) |
| 65-69 ans | 3,8 (5,84 %) | 40 520 [40 404 ; 40 635] | 42 476 | 1 841 (4,54 % [4,53 % ; 4,56 %]) | 988 (2,44 % [2,43 % ; 2,45 %]) |
| 70-74 ans | 3,4 (5,22 %) | 55 757 [54 681 ; 56 834] | 58 327 | 1 493 (2,68 % [2,63 % ; 2,73 %]) | 3 109 (5,58 % [5,47 % ; 5,69 %]) |
| 75-79 ans | 2,16 (3,32 %) | 51 767 [51 303 ; 52 232] | 57 516 | 5 284 (10,21 % [10,12 % ; 10,3 %]) | 2 606 (5,03 % [4,99 % ; 5,08 %]) |
| 80-84 ans | 1,83 (2,82 %) | 77 960 [76 645 ; 79 276] | 83 367 | 4 091 (5,25 % [5,16 % ; 5,34 %]) | 3 104 (3,98 % [3,92 % ; 4,05 %]) |
| 85-89 ans | 1,35 (2,08 %) | 116 220 [112 979 ; 119 462] | 121 025 | 1 563 (1,34 % [1,31 % ; 1,38 %]) | 1 124 (0,97 % [0,94 % ; 0,99 %]) |
| 90-94 ans | 0,66 (1,02 %) | 112 700 [110 672 ; 114 728] | 120 443 | 5 714 (5,07 % [4,98 % ; 5,16 %]) | 5 035 (4,47 % [4,39 % ; 4,55 %]) |
| 95-99 ans | 0,22 (0,34 %) | 81 354 [78 771 ; 83 937] | 78 384 | - 387 (-0,48 % [-0,46 % ; -0,49 %]) | 1 047 (1,29 % [1,25 % ; 1,33 %]) |
| Somme | 65,12 (100 %) | 629 490 [628 011 ; 630 970] | 654 016 | 23 046 (3,66 % [3,65 % ; 3,67 %]) | 16 309 (2,59 % [2,58 % ; 2,6 %]) |

Source : Insee, Indicateurs démographiques, calculs des auteurs.
*NB : L'excès (ou le défaut) de décès pour les 2 grands groupes d'âge (supérieur et inférieur à 65 ans) est la somme des excès (ou défaut) estimés au niveau de chaque âge pour prendre en considération l'évolution démographique de la manière la plus fine possible. Ceci explique la légère différence de résultat entre cette estimation et celle obtenue avec un calcul direct sur les 2 grands groupes d'âge (supérieur et inférieur à 65 ans).

Figure 1: Résultats Détaillés de l'Analyse des Auteurs

Dans l'ensemble, la surmortalité en 2020 par rapport aux trois dernières années (2017-2019) était entre 23 046 et 16 309 décès supplémentaires pour l'ensemble de la population.

L'augmentation de la mortalité due à l'épidémie de Covid-19 en 2020 a été estimée entre 3,66% et 2,59%, en fonction du mode de calcul utilisé par nos auteurs. Il est également intéressant de noter que la surmortalité a été particulièrement élevée chez les personnes âgées de plus de 65 ans, avec une surmortalité de 17 920 décès, soit **une augmentation de 3,34%**. En revanche, chez les personnes de moins de 65 ans, il n'y a pas eu de surmortalité notable en 2020. Ce sont ces chiffres qui nous ont particulièrement alerté car ils sont bien inférieurs à ceux avancés par l'INSEE autour de 9%.

Les auteurs concluent leur article en argumentant que l'impact de l'épidémie de Covid-19 sur la mortalité est moins important que ce qui a souvent été présenté dans le débat public, notamment en raison de la comparaison avec des maladies passées comme la grippe espagnole. Les auteurs notent que la mortalité en 2019 était exceptionnellement basse, et que la hausse observée en 2020 pourrait être en partie un rattrapage de cette sous-mortalité. En ce qui concerne les prédictions de mortalité dues à la Covid-19, les auteurs critiquent le recours à des modèles prédictifs qui ont anticipé un nombre très élevé de décès, arguant que ces prédictions n'ont pas été corroborées par les chiffres réels. Enfin, les auteurs mettent en garde contre une interprétation erronée de l'augmentation de la mortalité en 2020 comme étant entièrement due à la Covid-19. Ils soutiennent que cette augmentation est en partie due au vieillissement de la population et à la fin de vie des générations du baby-boom. En somme, ils soutiennent que la Covid-19 devrait être vue comme une autre des nombreuses maladies virales causant des pathologies respiratoires sévères n'en déplaise à toutes les institutions scientifiques sérieuses qui prônent le contraire.

2.2 Incohérences et questions potentielles

Nous avons rassemblé ici toutes les pistes et les évidences qui pourraient expliquer les potentielles erreurs ou négligences des auteurs.

Manipulation des intervalles de confiance : Tout d'abord, les auteurs jouent avec les bornes des intervalles de confiance pour minimiser l'impact du Covid-19. En effet, dans le tableau 1A (Figure 2), les auteurs débouchent sur 654 016 décès observés et 629 490 décès attendus. Si on les soustrait, on obtient 24 526 décès en excès, et non 23 046. Ce dernier chiffre a été trouvé en prenant la borne supérieure des décès attendus (630 970), ce qui nous permet d'arriver à leur résultat de 23 046. Cette méthode permet de réduire la surmortalité estimée en 2020.

Tableau 1a : Evaluation de la surmortalité en 2020 (3 années de référence : 2017-19)

| Classe d'âge | Population en million d'habitants | Décès attendus | Décès observés | Décès en excès sur classe d'âge | Décès en excès par âge |
|--------------|-----------------------------------|-----------------------------|----------------|-----------------------------------|-------------------------------------|
| 0-64 ans | 51,68 (79,36 %) | 93 211 [91 612 ; 94 810] | 92 478 | (0 % [0 % ; 0 %]) | - 705 (-0,76 % [-0,74 % ; -0,77 %]) |
| 65-99 ans | 13,44 (20,64 %) | 536 279 [528 940 ; 543 618] | 561 538 | 17 920 (3,34 % [3,3 % ; 3,39 %]) | 17 013 (3,17 % [3,13 % ; 3,22 %]) |
| Somme | 65,12 (100 %) | 629 490 [628 011 ; 630 970] | 654 016 | 23 046 (3,66 % [3,65 % ; 3,67 %]) | 16 309 (2,59 % [2,58 % ; 2,6 %]) |

Figure 2: Résultats Agrégés de l'Analyse des Auteurs

Écart entre les intervalles de confiance : Les bornes des intervalles de confiance des décès attendus nous ont assez surpris. En effet, les bornes pour les décès attendus en 2020 sont : [628 011 ; 630 970] alors qu'en 2019 c'est [617 657 ; 632 850]. L'écart des bornes est beaucoup plus grand en 2019 qu'en 2020. Ainsi, nous allons devoir creuser sur les intervalles de confiance des décès attendus en 2020 pour vérifier le travail des auteurs.

Opacité des résultats : Enfin, ils y a une certaine opacité dans leurs résultats. D'un côté ils ne donnent pas leurs bases de données, et de l'autre, ils n'explicitent pas le modèle statistiques utilisé pour estimer leurs quotients de mortalité.

2.3 A retenir

L'article que nous étudions ici présente une analyse qui, bien qu'elle se revendique précise et impartiale, laisse plusieurs zones d'ombre importantes. Les auteurs s'attardent sur des généralités statistiques connues de tous, mais négligent de donner des détails essentiels sur les aspects plus techniques de leur méthodologie. Cela inclut notamment leur manipulation des intervalles de confiance ainsi que l'opacité de leur modèle statistique et des bases de données utilisées qui constituent un aspect crucial de toute analyse statistique rigoureuse.

Résultats surprenants : Les auteurs aboutissent à des conclusions qui sont, au mieux, surprenantes. Ils trouvent une surmortalité notablement plus faible que ce que les chiffres officiels ont rapporté. Cela soulève des questions légitimes sur la validité de leurs résultats.

Approche méthodologique questionnable Leur approche est rapide et, à bien des égards, précaire. Ils sautent sur des concepts clés sans leur accorder le temps ou l'importance qu'ils méritent. Leur traitement des intervalles de confiance, par exemple, est particulièrement problématique, car il semble qu'ils utilisent ces outils de manière à minimiser l'impact du Covid-19 sur la mortalité en 2020.

Opacité des données et de la méthodologie : Ils font preuve d'une certaine opacité à la fois dans leur gestion des données traitées et dans leur méthodologie. Ils ne fournissent pas leurs bases de données et n'expliquent pas clairement la méthode démographique qu'ils utilisent pour estimer leurs quotients de mortalité. Cela rend difficile pour les lecteurs d'évaluer l'exactitude de leurs conclusions.

Objectif de notre analyse : Notre objectif, à travers cette analyse, est de démêler les fils de leur argumentation, de comprendre leurs méthodes et de voir si leurs conclusions tiennent réellement la route. En fin de compte, nous souhaitons déterminer si les auteurs ont raisonnablement minimisé l'impact du Covid-19 sur la mortalité en France en 2020, ou si, comme nous le soupçonnons, ils ont sous-estimé la véritable gravité de l'épidémie.

Cette étude est importante non seulement pour comprendre l'impact réel de la Covid-19 sur notre société, mais aussi pour souligner l'importance d'une analyse statistique rigoureuse et transparente. Dans une ère où l'information est omniprésente, il est plus crucial que jamais de savoir distinguer les analyses sérieuses des interprétations hâtives ou trompeuses.

3 Modèle statistique

3.1 Définition des variables

Afin d'établir une prédiction du nombre de décès attendus en France en 2020, nous devons réussir à prédire les quotients de mortalité en France en 2020 si la Covid-19 n'était jamais apparue. Ces quotients représentent la probabilité qu'un individu a de mourir pendant l'année selon son âge et son sexe. Nous devons donc définir l'ensemble des variables à l'oeuvre dans notre expérience.

3.1.1 Effectifs de population

Nous définissons :

- n_t l'effectif total de population pour une année t donnée.

3.1.2 Variables aléatoires

Pour toute année t , nous disposons d'un échantillon de n_t individus i.i.d mais hétérogènes. Pour tout individu i vivant à l'année t , nous définissons les variables aléatoires suivantes :

- $A_{i,t}$ qui représente l'âge de l'individu i au 31 décembre de l'année t .
- $S_{i,t}$ qui représente le sexe de l'individu i durant l'année t .
- $N_{t,a,s} = \sum_{i=1}^{n_t} (\mathbb{1}\{A_{i,t} = a\} \cdot \mathbb{1}\{S_{i,t} = s\})$ qui représente le nombre de personnes d'âge a et de sexe s durant l'année t .

Ainsi, en appliquant la **loi forte des grands nombres**, nous avons : $\forall t \in [t_{min}; t_{max}], \forall a \in [a_{min}; a_{max}], \forall s \in \{0, 1\}$

$$\begin{aligned} \frac{N_{t,a}}{n_t} &\xrightarrow[n_t \rightarrow \infty]{p.s} \mathbb{P}(A_{i,t} = a) \\ \frac{N_{t,s}}{n_t} &\xrightarrow[n_t \rightarrow \infty]{p.s} \mathbb{P}(S_{i,t} = s) \\ \frac{N_{t,a,s}}{n_t} &\xrightarrow[n_t \rightarrow \infty]{p.s} \mathbb{P}(A_{i,t} = a \cap S_{i,t} = s) = p_{t,a,s} \end{aligned}$$

- $X_{i,t}$ qui représente la probabilité pour l'individu i de mourir durant l'année t .

Nous avons :

$$X_{i,t} = \begin{cases} 1 & \text{si l'individu } i \text{ est mort durant l'année } t \\ 0 & \text{sinon} \end{cases}$$

Nous définissons alors

$$q_{i,t,a,s} = \mathbb{E}[X_{i,t} \mid A_{i,t} = a, S_{i,t} = s]$$

Nous avons donc $((X_{i,t} \mid A_{i,t} = a, S_{i,t} = s))_{i=1,\dots,n_t} \stackrel{iid}{\sim} \mathcal{B}(q_{t,a,s})$ une suite de variables aléatoires indépendantes suivant une même loi de Bernoulli de paramètre $q_{t,a,s}$. De façon intuitive, nous modélisons notre expérience comme une urne de Bernoulli avec une probabilité conditionnelle de succès.

Exemple :

$q_{2020,30,1} = \mathbb{E}[X_{1,2020} \mid A_{1,2020} = 30, S_{1,2020} = 1]$ représente la probabilité pour un individu de mourir au cours de l'année 2020, sachant que c'est un homme de 30 ans .

3.1.3 Quotients de mortalité observés

Afin de prédire nos quotients de mortalité en 2020, nous allons utiliser ceux des années précédentes. Ces quotients que nous pouvons observé dans les tables de l'INSEE sont des réalisations des variables aléatoires introduites précédemment. Nous pouvons définir ces quotients de mortalité observés $\hat{q}_{t,a,s}$ à année t , âge a et sexe s donnés :

$$\hat{q}_{t,a,s} = \frac{1}{N_{t,a,s}} \cdot \sum_{i=1}^{n_t} (X_{i,t} \cdot \mathbb{1}\{A_{i,t} = a\} \cdot \mathbb{1}\{S_{i,t} = s\})$$

avec $\mathbb{E}[X_{i,t} \mid A_{i,t} = a, S_{i,t} = s] = q_{t,a,s}$ et $\mathbb{V}[X_{i,t} \mid A_{i,t} = a, S_{i,t} = s] = q_{t,a,s} \cdot (1 - q_{t,a,s})$.

Pour toute période $\mathcal{T} = [\underline{t}; \bar{t}]$ (les auteurs s'intéressent notamment aux années 2017, 2018 et 2019 dans leur article pour prédire les quotients de 2020), nous définissons le vecteur $\hat{q}_{\mathcal{T},a,s}$ des quotients de mortalité à période \mathcal{T} , âge a et sexe s donnés:

$$\hat{q}_{\mathcal{T},a,s} = \begin{pmatrix} \hat{q}_{\underline{t},a,s} \\ \dots \\ \hat{q}_{\bar{t},a,s} \end{pmatrix} = \begin{pmatrix} \frac{1}{N_{\underline{t},a,s}} \sum_{i=1}^{n_{\underline{t}}} (X_{i,\underline{t}} \cdot \mathbb{1}\{A_{i,\underline{t}} = a\} \cdot \mathbb{1}\{S_{i,\underline{t}} = s\}) \\ \dots \\ \frac{1}{N_{\bar{t},a,s}} \sum_{i=1}^{n_{\bar{t}}} (X_{i,\bar{t}} \cdot \mathbb{1}\{A_{i,\bar{t}} = a\} \cdot \mathbb{1}\{S_{i,\bar{t}} = s\}) \end{pmatrix}$$

Exemple :

En considérant $\mathcal{T} = \llbracket 2017; 2019 \rrbracket$, $a = 30$ et $s = 1$ nous avons :

$$\hat{q}_{\llbracket 2017; 2019 \rrbracket, 30, 1} = \begin{pmatrix} \hat{q}_{2017, 30, 1} \\ \dots \\ \hat{q}_{2019, 30, 1} \end{pmatrix} = \begin{pmatrix} \frac{1}{N_{2017, 30, 1}} \sum_{i=1}^{n_{2017}} (X_{i, 2017} \cdot \mathbb{1}\{A_{i, 2017} = 30\} \cdot \mathbb{1}\{S_{i, 2017} = 1\}) \\ \dots \\ \frac{1}{N_{2019, 30, 1}} \sum_{i=1}^{n_{2019}} (X_{i, 2019} \cdot \mathbb{1}\{A_{i, 2019} = 30\} \cdot \mathbb{1}\{S_{i, 2019} = 1\}) \end{pmatrix}$$

qui représente le vecteur des quotients de mortalité de 2017 à 2019 pour les hommes de 30 ans, et qui nous permettra ensuite d'estimer le quotient de mortalité de 2020 pour les hommes de 30 ans.

3.1.4 Quotients de mortalité prédits

Une fois les quotients de mortalité passés définis, il nous faut créer l'estimateur de nos paramètres d'intérêts, à savoir: les quotients de mortalité en 2020. Pour cela, nous définissons f , une fonction différentiable :

$$\begin{aligned} f &: \mathbb{R}^d && \rightarrow \mathbb{R}^p \\ (x_1, \dots, x_d) &\mapsto f(x_1, \dots, x_d) \end{aligned}$$

Nous pouvons alors définir $\tilde{q}_{\tilde{t}, a, s}$ une nouvelle variable visant à prédire les quotients de mortalité pour l'année \tilde{t} , à âge a et sexe s donnés en fonction des quotients de mortalité passés :

$$\tilde{q}_{\tilde{t}, a, s} = f(\hat{q}_{\mathcal{T}, a, s})$$

Cette fonction f a pour but de prédire un quotient de mortalité en utilisant les quotients de mortalité passés. Nous pouvons citer quelques exemples :

- définir le quotient de mortalité comme celui de l'année précédente
- définir le quotient de mortalité comme la moyenne des quotients des trois années précédentes. C'est notamment la méthode retenue par nos auteurs.

Nous posons également $q_{\tilde{t}, a, s}^P = f(q_{\mathcal{T}, a, s})$. Cette variable théorique représente la prédiction théorique faisant appel à notre fonction f . Elle diffère de $\tilde{q}_{\tilde{t}, a, s}$ dans le sens où $q_{\tilde{t}, a, s}^P$ ne repose pas sur la réalisation de variables aléatoires. Nous ferons appel à cette prédiction théorique dans la création de nos intervalles de confiance.

Exemple :

En considérant $\mathcal{T} = \llbracket 2017; 2019 \rrbracket$, $a = 30$, $s = 1$ et $\tilde{t} = 2020$, nous avons :

$$\tilde{q}_{2020,30,1} = f(\hat{q}_{\llbracket 2017; 2019 \rrbracket, 30, 1})$$

qui représente le vecteur des quotients de mortalité prédits pour les hommes de 30 ans en 2020 en fonction des quotients de mortalité observés pour les hommes de 30 ans en 2017, 2018 et 2019.

3.2 Quotients de mortalité observés : théorèmes limites

Afin de pouvoir construire nos intervalles de confiance ou de conduire des tests sur notre prédiction, nous avons besoin d'étudier la convergence de nos estimateurs des quotients de mortalités observés. Nous rappelons que nous avons :

$$N_{t,a,s} = \sum_{i=1}^{n_t} (\mathbb{1}\{A_{i,t} = a\} \cdot \mathbb{1}\{S_{i,t} = s\})$$

et

$$\hat{q}_{t,a,s} = \frac{1}{N_{t,a,s}} \sum_{i=1}^{n_t} (X_{i,t} \cdot \mathbb{1}\{A_{i,t} = a\} \cdot \mathbb{1}\{S_{i,t} = s\}) = \left(\frac{n_t}{N_{t,a,s}} \right) \frac{1}{n_t} \sum_{i=1}^{n_t} (X_{i,t} \cdot \mathbb{1}\{A_{i,t} = a\} \cdot \mathbb{1}\{S_{i,t} = s\})$$

avec pour tout i (nous rappelons que nous avons $p_{t,a,s} = \mathbb{P}(A_t = a \cap S_t = s)$) :

- $\mathbb{E}[X_{i,t} \mid A_{i,t} = a, S_{i,t} = s] = q_{t,a,s}$
- $\mathbb{V}[X_{i,t} \mid A_{i,t} = a, S_{i,t} = s] = q_{t,a,s} \cdot (1 - q_{t,a,s})$
- $\mathbb{E}[X_{i,t} \cdot \mathbb{1}\{A_{i,t} = a\} \cdot \mathbb{1}\{S_{i,t} = s\}] = p_{t,a,s} \cdot q_{t,a,s}$
- $\mathbb{V}[X_{i,t} \cdot \mathbb{1}\{A_{i,t} = a\} \cdot \mathbb{1}\{S_{i,t} = s\}] = p_{t,a,s} \cdot q_{t,a,s} \cdot (1 - p_{t,a,s} \cdot q_{t,a,s})$

D'où, d'après la **loi forte des grands nombres**, nous avons :

$$\frac{N_{t,a,s}}{n_t} \xrightarrow[n_t \rightarrow \infty]{p.s.} p_{t,a,s}$$

et

$$\frac{1}{n_t} \sum_{i=1}^{n_t} (X_{i,t} \cdot \mathbb{1}\{A_{i,t} = a\} \cdot \mathbb{1}\{S_{i,t} = s\}) \xrightarrow[n_t \rightarrow \infty]{p.s.} p_{t,a,s} \cdot q_{t,a,s}$$

D'où, en appliquant le **premier théorème de continuité** à la fonction $(x, y) \mapsto \frac{x}{y}$ avec

$y > 0$, nous avons :

$$\hat{q}_{t,a,s} \xrightarrow[n_t \rightarrow \infty]{p.s} q_{t,a,s}$$

Nous pouvons également appliquer le **théorème central limite**, et nous obtenons donc :

$$\sqrt{n_t} \left(\begin{array}{c} \frac{1}{n_t} \sum_{i=1}^{n_t} (X_{i,t} \cdot \mathbb{1}\{A_{i,t} = a\} \cdot \mathbb{1}\{S_{i,t} = s\}) - p_{t,a,s} \cdot q_{t,a,s} \\ \frac{1}{n_t} \sum_{i=1}^{n_t} (\mathbb{1}\{A_{i,t} = a\} \cdot \mathbb{1}\{S_{i,t} = s\}) - p_{t,a,s} \end{array} \right) \xrightarrow[n_t \rightarrow \infty]{d} \mathcal{N}(0, \Sigma_{t,a,s})$$

où

$$\Sigma_{t,a,s} = \begin{pmatrix} \mathbb{V}[X_t \cdot \mathbb{1}\{A_t = a\} \cdot \mathbb{1}\{S_t = s\}] & c \\ c & \mathbb{V}[\mathbb{1}\{A_t = a\} \cdot \mathbb{1}\{S_t = s\}] \end{pmatrix}$$

avec

$$c = \text{Cov}(X_t \cdot \mathbb{1}\{A_t = a\} \cdot \mathbb{1}\{S_t = s\}, \mathbb{1}\{A_t = a\} \cdot \mathbb{1}\{S_t = s\})$$

Nous pouvons alors appliquer la **delta-méthode** à la fonction suivante :

$$\begin{array}{lll} h & : & \mathbb{R} \times \mathbb{R}^* \rightarrow \mathbb{R} \\ & & (x, y) \mapsto \frac{x}{y} \end{array}$$

qui est bien différentiable pour $y > 0$ et telle que $\nabla h(x, y) = (\frac{1}{y}, \frac{-x}{y^2})$

Et nous obtenons ainsi, après calculs :

$$\sqrt{n_t} (\hat{q}_{t,a,s} - q_{t,a,s}) \xrightarrow[n_t \rightarrow \infty]{d} \mathcal{N}\left(0, \frac{q_{t,a,s} \cdot (1 - q_{t,a,s})}{p_{t,a,s}}\right)$$

i.e.

$$\sqrt{n_t \cdot p_{t,a,s}} \frac{\hat{q}_{t,a,s} - q_{t,a,s}}{\sqrt{q_{t,a,s} \cdot (1 - q_{t,a,s})}} \xrightarrow[n_t \rightarrow \infty]{d} \mathcal{N}(0, 1)$$

Et donc, d'après **théorème de Slutsky**, nous avons :

$$\sqrt{N_{t,a,s}} \frac{\hat{q}_{t,a,s} - q_{t,a,s}}{\sqrt{\hat{q}_{t,a,s} \cdot (1 - \hat{q}_{t,a,s})}} \xrightarrow[n_t \rightarrow \infty]{d} \mathcal{N}(0, 1)$$

Pour passer à l'analyse multi-variée nous ferons l'hypothèse suivante :

$$\forall t \in \mathcal{T}, \frac{n_t}{n_t} \xrightarrow[n_t, n_{\underline{t}} \rightarrow +\infty]{} \lambda_t \in \mathbb{R}$$

Cela revient à dire que les tailles d'échantillons croissent à la même vitesse.

De la même manière que précédemment, nous pouvons étendre notre raisonnement à un vecteur de quotients de mortalité.

En effet, d'après la **loi forte des grands nombres** et le **théorème central limite** nous avons :

$$\hat{q}_{\mathcal{T},a,s} \xrightarrow[n_t \rightarrow \infty]{p.s} q_{\mathcal{T},a,s} = \begin{pmatrix} q_{\underline{t},a,s} \\ \dots \\ q_{\bar{t},a,s} \end{pmatrix}$$

et

$$\sqrt{n_{\underline{t}}} \cdot (\hat{q}_{\mathcal{T},a,s} - q_{\mathcal{T},a,s}) \xrightarrow[n_t \rightarrow \infty]{d} \mathcal{N}(0, \Sigma_{\mathcal{T},a,s})$$

où

$$\Sigma_{\mathcal{T},a,s} = \begin{bmatrix} \frac{q_{\underline{t},a,s} \cdot (1 - q_{\underline{t},a,s}) \cdot \lambda_{\underline{t}}}{p_{\underline{t},a,s}} & & (0) \\ & \ddots & \\ (0) & & \frac{q_{\bar{t},a,s} \cdot (1 - q_{\bar{t},a,s}) \cdot \lambda_{\bar{t}}}{p_{\bar{t},a,s}} \end{bmatrix}$$

$$\hat{\Sigma}_{\mathcal{T},a,s} = \begin{bmatrix} \frac{\hat{q}_{\underline{t},a,s} \cdot (1 - \hat{q}_{\underline{t},a,s}) \cdot n_{\underline{t}}}{N_{\underline{t},a,s}} & & (0) \\ & \ddots & \\ (0) & & \frac{\hat{q}_{\bar{t},a,s} \cdot (1 - \hat{q}_{\bar{t},a,s}) \cdot n_{\bar{t}}}{N_{\bar{t},a,s}} \end{bmatrix}$$

Nous définissons alors : $\hat{\Sigma}_{\mathcal{T},a,s}$

$$\hat{\Sigma}_{\mathcal{T},a,s} \xrightarrow[n_t \rightarrow \infty]{p.s} \Sigma_{\mathcal{T},a,s}$$

Et ainsi, d'après le **théorème de Slutsky** :

$$\sqrt{n_{\underline{t}}} \frac{\hat{q}_{\mathcal{T},a,s} - q_{\mathcal{T},a,s}}{\sqrt{\hat{\Sigma}_{\mathcal{T},a,s}}} \xrightarrow[n_t \rightarrow \infty]{d} \mathcal{N}(0, 1)$$

3.3 Quotients de mortalité prédits : théorèmes limites

Maintenant que nous avons établi la convergence de nos quotients observés, nous allons pouvoir passer à l'étude de la convergence de nos quotients prédits. Ces derniers sont fonctions des quotients observés et nous allons donc pouvoir établir les convergences nécessaires

pour nos analyses futures. Nous rappelons que nous avons :

$$\hat{q}_{\mathcal{T},a,s} \xrightarrow[n_t \rightarrow \infty]{p.s} q_{\mathcal{T},a,s} = \begin{pmatrix} q_{t,a,s} \\ \dots \\ q_{\bar{t},a,s} \end{pmatrix}$$

et

$$\sqrt{n_{\underline{t}}} \cdot (\hat{q}_{\mathcal{T},a,s} - q_{\mathcal{T},a,s}) \xrightarrow[n_t \rightarrow \infty]{d} \mathcal{N}(0, \Sigma_{\mathcal{T},a,s})$$

où

$$\Sigma_{\mathcal{T},a,s} = \begin{bmatrix} \frac{q_{t,a,s} \cdot (1 - q_{t,a,s}) \cdot \lambda_t}{p_{t,a,s}} & & (0) \\ & \ddots & \\ (0) & & \frac{q_{\bar{t},a,s} \cdot (1 - q_{\bar{t},a,s}) \cdot \lambda_{\bar{t}}}{p_{\bar{t},a,s}} \end{bmatrix}$$

Ainsi, grâce au **premier théorème de continuité** et à la **delta-méthode** appliqués à la fonction f qui est bien continue et différentiable (par hypothèse), nous avons :

$$f(\hat{q}_{\mathcal{T},a,s}) \xrightarrow[n_t \rightarrow \infty]{p.s} f(q_{\mathcal{T},a,s})$$

et

$$\sqrt{n_{\underline{t}}} (f(\hat{q}_{\mathcal{T},a,s}) - f(q_{\mathcal{T},a,s})) \xrightarrow[n_t \rightarrow \infty]{d} \mathcal{N}(0, \nabla f(q_{\mathcal{T},a,s})^T \Sigma_{\mathcal{T},a,s} \nabla f(q_{\mathcal{T},a,s}))$$

i.e.

$$\tilde{q}_{\bar{t},a,s} \xrightarrow[n_t \rightarrow \infty]{p.s} q_{\bar{t},a,s}^P \text{ (convergence vers un quotient prédit)}$$

et

$$\sqrt{n_{\underline{t}}} (\tilde{q}_{\bar{t},a,s} - q_{\bar{t},a,s}^P) \xrightarrow[n_t \rightarrow \infty]{d} \mathcal{N}(0, \nabla f(q_{\mathcal{T},a,s})^T \Sigma_{\mathcal{T},a,s} \nabla f(q_{\mathcal{T},a,s}))$$

Nous avons alors, d'après le **premier théorème de continuité** et le **théorème de Slutsky** :

$$\frac{\sqrt{n_{\underline{t}}} (\tilde{q}_{\bar{t},a,s} - q_{\bar{t},a,s}^P)}{\sqrt{\frac{1}{\lambda_{\bar{t}}} \nabla f(\hat{q}_{\mathcal{T},a,s})^T \hat{\Sigma}_{\mathcal{T},a,s} \nabla f(\hat{q}_{\mathcal{T},a,s})}} \xrightarrow[n_t \rightarrow \infty]{d} \mathcal{N}(0, 1)$$

3.4 Intervalles de confiance

Dans cette sous-partie, nous allons déterminer des intervalles de confiance au niveau α des quotients de mortalité observés et prédits au cours d'une année t .

3.4.1 Quotients de mortalité observés

Nous commençons par établir les intervalles de confiance pour nos quotients de mortalité observés. Nous rappelons que nous avons :

$$\sqrt{n_{\tilde{t}}} \frac{\hat{q}_{\tilde{t},a,s} - q_{\tilde{t},a,s}}{\sqrt{\hat{\Sigma}_{\tilde{t},a,s}}} \xrightarrow[n_{\tilde{t} \rightarrow \infty}{d} \mathcal{N}(0, 1)$$

Soit $(b, c) \in \llbracket 0; 1 \rrbracket^2$, nous avons :

$$\mathbb{P} \left(\sqrt{n_{\tilde{t}}} \frac{\hat{q}_{\tilde{t},a,s} - q_{\tilde{t},a,s}}{\sqrt{\hat{\Sigma}_{\tilde{t},a,s}}} \in [b, c] \right) \xrightarrow[n_t \rightarrow \infty]{p.s} \mathbb{P}(\mathcal{N}(0, 1) \in [b, c])$$

D'où

$$\mathbb{P} \left(\sqrt{n_{\tilde{t}}} \frac{\hat{q}_{\tilde{t},a,s} - q_{\tilde{t},a,s}}{\sqrt{\hat{\Sigma}_{\tilde{t},a,s}}} \in [-z_{1-\alpha/2}; z_{1-\alpha/2}] \right) \xrightarrow[n_t \rightarrow \infty]{p.s} 1 - \alpha$$

où z_r représente le r -ième quantile d'une loi normale centrée réduite.

Nous avons donc :

$$\mathbb{P} \left(q_{\tilde{t},a,s} \in \left[\hat{q}_{\tilde{t},a,s} \pm z_{1-\alpha/2} \sqrt{\frac{\hat{\Sigma}_{\tilde{t},a,s}}{n_{\tilde{t}}}} \right] \right) \xrightarrow[n_t \rightarrow \infty]{p.s} 1 - \alpha$$

Nous pouvons alors multiplier les bornes de l'intervalle de confiance par l'effectif de population associé pour déterminer un intervalle de confiance de niveau α du nombre de décès observés durant l'année \tilde{t} à âge et sexe donnés.

Désormais, nous souhaitons déterminer un intervalle de confiance de niveau α pour le quotient de mortalité moyen sur l'ensemble de la population durant l'année \tilde{t} .

Nous posons alors :

$$\hat{q}_{\tilde{t}} = \sum_{a,s} \frac{N_{\tilde{t},a,s}}{n_{\tilde{t}}} \hat{q}_{\tilde{t},a,s}$$

Nous avons :

$$\hat{q}_{\tilde{t}} \xrightarrow[n_t \rightarrow \infty]{p.s} \sum_{a,s} p_{t,a,s} q_{\tilde{t},a,s} = q_{\tilde{t}}$$

et donc, par indépendance des $(\hat{q}_{\tilde{t},a,s})_{a,s}$,

$$\mathbb{P} \left(q_{\tilde{t}} \in \left[\hat{q}_{\tilde{t}} \pm z_{1-\alpha/2} \sum_{a,s} \frac{N_{\tilde{t},a,s}}{n_{\tilde{t}}} \sqrt{\frac{\hat{\Sigma}_{\tilde{t},a,s}}{n_{\tilde{t}}}} \right] \right) \xrightarrow[n_t \rightarrow \infty]{p.s} 1 - \alpha$$

Nous pouvons alors multiplier les bornes de l'intervalle de confiance par l'effectif de population associé pour déterminer un intervalle de confiance de niveau α du nombre de décès observés durant l'année \tilde{t} .

3.4.2 Quotients de mortalité prédits

Maintenant, nous pouvons passer à la construction de nos intervalles de confiance pour les quotients prédits. Nous rappelons que nous avons :

$$\frac{\tilde{q}_{\tilde{t},a,s} - q_{\tilde{t},a,s}^P}{\sqrt{\frac{1}{\lambda_{\tilde{t}}} \nabla f(\hat{q}_{\mathcal{T},a,s})^T \hat{\Sigma}_{\mathcal{T},a,s} \nabla f(\hat{q}_{\mathcal{T},a,s})}} \xrightarrow[n_t \rightarrow \infty]{d} \mathcal{N}(0, 1)$$

Ainsi, de la même manière que précédemment, nous avons :

$$\mathbb{P} \left(q_{\tilde{t},a,s}^P \in \left[\tilde{q}_{\tilde{t},a,s} \pm z_{1-\alpha/2} \sqrt{\frac{\frac{1}{\lambda_{\tilde{t}}} \nabla f(\hat{q}_{\mathcal{T},a,s})^T \hat{\Sigma}_{\mathcal{T},a,s} \nabla f(\hat{q}_{\mathcal{T},a,s})}{n_{\tilde{t}}}} \right] \right) \xrightarrow[n_t \rightarrow \infty]{p.s} 1 - \alpha$$

Nous pouvons alors multiplier les bornes de l'intervalle de confiance par l'effectif de population associé pour déterminer un intervalle de confiance de niveau α du nombre de décès observés durant l'année \tilde{t} à âge et sexe donnés.

De même, en reprenant la même méthode que précédemment, nous avons :

$$\mathbb{P} \left(q_{\tilde{t}}^P \in \left[\hat{q}_{\tilde{t}} \pm z_{1-\alpha/2} \sum_{a,s} \frac{N_{\tilde{t},a,s}}{n_{\tilde{t}}} \sqrt{\frac{\frac{1}{\lambda_{\tilde{t}}} \nabla f(\hat{q}_{\mathcal{T},a,s})^T \hat{\Sigma}_{\mathcal{T},a,s} \nabla f(\hat{q}_{\mathcal{T},a,s})}{n_{\tilde{t}}}} \right] \right) \xrightarrow[n_t \rightarrow \infty]{p.s} 1 - \alpha$$

Nous pouvons alors multiplier les bornes de l'intervalle de confiance par l'effectif de population associé pour déterminer un intervalle de confiance de niveau α du nombre de décès observés durant l'année \tilde{t} .

3.5 Tests de surmortalité

3.5.1 Test de surmortalité à âge et sexe donnés

Dans cette sous-partie, nous souhaitons déterminer s'il y a eu une différence significative de la mortalité en \tilde{t} , à âge et sexe donnés.

Nous devons donc tester :

$$H0 : q_{\tilde{t},a,s} = q_{\tilde{t},a,s}^P \text{ contre } H1 : q_{\tilde{t},a,s} \neq q_{\tilde{t},a,s}^P$$

Pour cela, nous allons déterminer un intervalle de confiance de niveau α de la quantité $q_{\tilde{t},a,s} - q_{\tilde{t},a,s}^P$ pour déterminer sa significativité et la quantifier.

Nous avons pour cela un estimateur consistant :

$$\hat{q}_{\tilde{t},a,s} - \tilde{q}_{\tilde{t},a,s} \xrightarrow[n_t \rightarrow \infty]{p.s} q_{\tilde{t},a,s} - q_{\tilde{t},a,s}^P$$

De plus, $\hat{q}_{\tilde{t},a,s}$ et $\tilde{q}_{\tilde{t},a,s}$ sont asymptotiquement indépendants puisque nous n'utilisons pas les mêmes données pour les déterminer, et que ces données sont indépendantes.

En effet, $\hat{q}_{\tilde{t},a,s}$ utilise les données de \tilde{t} , tandis que $\tilde{q}_{\tilde{t},a,s}$ utilise les données des années précédentes, et que ces données sont supposées indépendantes.

Nous avons alors :

$$\mathbb{P} \left(q_{\tilde{t},a,s} - q_{\tilde{t},a,s}^P \in \left[\hat{q}_{\tilde{t},a,s} - \tilde{q}_{\tilde{t},a,s} \pm z_{1-\alpha/2} \sqrt{\frac{\frac{1}{\lambda_{\tilde{t}}} \nabla f(\hat{q}_{\mathcal{T},a,s})^T \hat{\Sigma}_{\mathcal{T},a,s} \nabla f(\hat{q}_{\mathcal{T},a,s}) + \hat{\Sigma}_{\tilde{T},a,s}}{n_{\tilde{t}}}} \right] \right) \xrightarrow[n_t \rightarrow \infty]{p.s} 1 - \alpha$$

Nous pouvons alors déterminer si l'écart entre $q_{\tilde{t},a,s}$ et $q_{\tilde{t},a,s}^P$ est significatif ou non au niveau α . Pour cela, il suffit de regarder si 0 appartient à l'intervalle de confiance déterminé précédemment. Si ce n'est pas le cas, alors l'écart entre les deux est significatif.

Pour quantifier cette différence, nous pouvons alors multiplier les bornes de l'intervalle de confiance par l'effectif de population associé à l'âge et le sexe.

3.5.2 Test de surmortalité global

Dans cette sous-partie, nous souhaitons déterminer s'il y a eu une différence significative de mortalité au cours de l'année \tilde{t} sur l'ensemble de la population.

Nous devons donc tester :

$$H_0 : q_{\tilde{t}} = q_{\tilde{t}}^P \text{ contre } H_1 : q_{\tilde{t}} \neq q_{\tilde{t}}^P$$

Pour cela, nous allons déterminer un intervalle de confiance de niveau α de la quantité $q_{\tilde{t}} - q_{\tilde{t}}^P$ pour déterminer sa significativité et la quantifier.

En raisonnant de la même manière que précédemment, nous obtenons :

$$\hat{q}_{\tilde{t}} - \tilde{q}_{\tilde{t}} \xrightarrow[n_{\tilde{t}} \rightarrow \infty]{p.s} q_{\tilde{t}} - q_{\tilde{t}}^P$$

et

$$\mathbb{P} \left(q_{\tilde{t}} - q_{\tilde{t}}^P \in \left[\hat{q}_{\tilde{t}} - \tilde{q}_{\tilde{t}} \pm z_{1-\alpha/2} \sum_{a,s} \frac{N_{\tilde{t},a,s}}{n_{\tilde{t}}} \sqrt{\frac{\frac{1}{\lambda_{\tilde{t}}} \nabla f(\hat{q}_{\mathcal{T},a,s})^T \hat{\Sigma}_{\mathcal{T},a,s} \nabla f(\hat{q}_{\mathcal{T},a,s}) + \hat{\Sigma}_{\tilde{T},a,s}}{n_{\tilde{t}}}} \right] \right) \xrightarrow[n_{\tilde{t}} \rightarrow \infty]{p.s} 1 - \alpha$$

Nous pouvons alors déterminer si l'écart entre $q_{\tilde{t}}$ et $q_{\tilde{t}}^P$ est significatif ou non au niveau α . Pour cela, il suffit de regarder si 0 appartient à l'intervalle de confiance déterminé précédemment. Si ce n'est pas le cas, alors l'écart entre les deux est significatif.

Pour quantifier cette différence, nous pouvons alors multiplier les bornes de l'intervalle de confiance par l'effectif de population associé.

4 Application du modèle : Résultats

Maintenant que nous avons établi un cadre théorique solide pour modéliser notre expérience, nous pouvons passer à la mise en pratique du modèle que nous venons de définir.

4.1 Mise en pratique du modèle

Dans un souci de rigueur scientifique, nous avons laissé en accès libre et documenté scrupuleusement chaque étape de notre mise en application.

Pour réaliser nos prédictions, nous utilisons 2 bases de données disponibles sur le site de l'INSEE :

- le T2 qui regroupe, pour 100 000 personnes de sexe et d'année de naissance donnés, le quotient de mortalité associé à chaque âge révolu.
- le T69QMORT qui regroupe, pour 100 000 personnes de sexe et d'âge donnés, le quotient de mortalité associé à chaque année..

Après quelques modifications, nous obtenons alors plusieurs matrices (représentées par des DataFrames) contenant nos $\hat{q}_{t,a,s}$.

Nous pouvons alors définir nos différentes fonctions f que nous utilisons pour obtenir nos $\tilde{q}_{t,a,s}$:

- f_1 qui définit les quotients de l'année 2020 comme étant ceux de l'année 2019. Cette approche basique permet à la fois de faciliter l'application de notre modèle et d'avoir une première idée du nombre de décès attendus en France en 2020. En effet, les quotients de mortalité étant relativement stables d'une année à une autre, il est courant de reprendre ceux d'une année précédente pour estimer le nombre de décès de l'année qui suit.
- f_2 qui définit les quotients de l'année 2020 comme étant la moyenne des quotients de 2017, 2018 et 2019. C'est la méthode retenue par les auteurs de notre article. Ce choix se justifie par la relative stabilité des quotients de mortalité ainsi que d'une tendance récente à la diminution du rythme de croissance des quotients de mortalité. En effet, les quotients de mortalité pour un sexe et un âge donné augmentent moins vite les 5 dernières années que depuis le début des années 2000. Il est donc raisonnable de s'arrêter aux 3 dernières années et de ne pas prendre en compte les précédents.

Pour chacune des bases de données, et pour chacune des fonctions, nous créons alors plusieurs matrices contenant nos $\tilde{q}_{t,a,s}$. Nous en profitons pour créer également des matrices associées contenant les bornes des intervalles de confiance des $\tilde{q}_{t,a,s}$.

Nous pouvons alors déterminer le nombre de décès prédits et les intervalles de confiance associés, que nous regroupons dans un tableau récapitulatif semblable à celui réalisé par les auteurs de l'article. Pour rappel, voici le tableau final des auteurs:

Tableau 1a : Evaluation de la surmortalité en 2020 (3 années de référence : 2017-19)

| Classe d'âge | Population en million d'habitants | Décès attendus | Décès observés | Décès en excès sur classe d'âge | Décès en excès par âge |
|--------------|-----------------------------------|-----------------------------|----------------|-----------------------------------|-------------------------------------|
| 0-64 ans | 51,68 (79,36 %) | 93 211 [91 612 ; 94 810] | 92 478 | (0 % [0 % ; 0 %]) | - 705 (-0,76 % [-0,74 % ; -0,77 %]) |
| 65-99 ans | 13,44 (20,64 %) | 536 279 [528 940 ; 543 618] | 561 538 | 17 920 (3,34 % [3,3 % ; 3,39 %]) | 17 013 (3,17 % [3,13 % ; 3,22 %]) |
| Somme | 65,12 (100 %) | 629 490 [628 011 ; 630 970] | 654 016 | 23 046 (3,66 % [3,65 % ; 3,67 %]) | 16 309 (2,59 % [2,58 % ; 2,6 %]) |

Figure 3 : Résultats de l'article

Et voilà les résultats que nous obtenons :

| | Population | Décès attendus | Décès observés | Différence de décès | Pourcentage de différence |
|-----------|------------|------------------|----------------|---------------------|---------------------------|
| (0, 64] | 51821513 | [93716, 100149] | 92478 | [-7671, -1238] | [-7.66 %, -1.32 %] |
| (64, 100] | 13443795 | [484223, 496919] | 561538 | [64619, 77315] | [13.0 %, 15.97 %] |
| Total | 65265308 | [577939, 597068] | 654016 | [56948, 76077] | [9.54 %, 13.16 %] |

Figure 4 : Résultats pour la première base de données et la fonction f_2

4.2 Résultats et interprétation

Nous pouvons désormais commenter nos résultats et les comparer avec ceux obtenus par les auteurs de l'article. Tous nos résultats sont disponibles en annexe.

Dans nos 4 applications, nous prédisons une surmortalité durant l'année 2020. En effet, nous prédisons entre 50 000 et 100 000 décès supplémentaires par rapport à ce qui était attendu en 2020. Nous pouvons constater que la seconde base de données semble sous-estimer les quotients de mortalité, tant nos prédictions semblent peu élevées (entre 529 575 et 561 188 décès attendus pour la fonction f_1 , alors qu'il y avait eu plus de 600 000 décès en 2019). Nous prendrons donc en considération les résultats issus de la première base de données, qui semblent bien plus réalistes.

En effet, la prédiction réalisée grâce à f_2 nous donne un intervalle de confiance à 95% du nombre de décès attendus égal à [577 939, 597 068], ce qui signifie **qu'il y a eu une surmortalité entre 9,54% et 13,16%**, ce qui se rapproche de la surmortalité annoncée par l'INSEE qui était de 9%. Notre estimation vient donc s'opposer encore plus clairement aux résultats des auteurs que celles de l'INSEE.

Néanmoins, il est intéressant de constater qu'il existe une forte dichotomie entre les moins de 65 ans et les plus de 65 ans. En effet, selon la prédiction f_1 , il n'y a pas eu de surmortalité chez les moins de 65 ans (puisque 0 appartient à l'intervalle de confiance à 95% de la différence du nombre de décès attendus et du nombre de décès observés). Mais la prédiction f_2 va même plus loin, puisque selon elle, il y a même eu une sous-mortalité chez les moins de 65 ans, avec entre 1 238 et 7 671 décès de moins que ce qui était attendu. Plusieurs

facteurs peuvent expliquer ces résultats : les différents confinements, le renforcement des gestes barrières, de meilleures pratiques sanitaires... Pour les plus de 65 ans en revanche, la surmortalité est bien plus élevée (**plus de 10%**), ce qui entraîne donc la surmortalité globale à la hausse. Il semble donc y avoir un consensus sur ce point : que ce soit les auteurs, l'INSEE ou nous, nous retrouvons tous ce phénomène de dichotomie en fonction de l'âge. Ce n'est pas simplement que les personnes âgées sont plus sévèrement touchées par les maladies infectieuses comme la Covid-19 (ce qui est complètement évident), cela signifie que la Covid-19 n'a pas engendré de surmortalité chez les personnes jeunes.

Pour résumer, si nous comparons nos résultats obtenus avec la première base de donnée et la fonction f_2 (pour reprendre la méthode utilisée dans l'article) avec ceux des auteurs de l'article, nous voyons des différences claires. Bien que, nous obtenions des résultats plutôt similaires en ce qui concerne les moins de 65 ans : il n'y a pas eu de surmortalité chez les moins de 65 ans en 2020, nos résultats diffèrent lourdement pour les plus de 65 ans. Comme les auteurs de l'article, nous obtenons une surmortalité chez les plus de 65 ans, mais notre surmortalité est bien plus élevée (**plus de 13% pour nous contre 3,17% dans l'article, soit près de 5 fois plus**). Cela entraîne donc une différence pour la surmortalité globale, encore une fois bien plus élevée selon nous (**plus de 9% selon nous contre 2,59% d'après l'article**).

Ainsi, nous sommes dans l'obligation de réfuter les conclusions avancées par les auteurs de l'article. Non, la Covid-19 n'a pas été qu'un simple épisode grippal, car elle a provoqué une surmortalité très importante, notamment chez les personnes de plus de 65 ans. Nous rejoignons et soutenons ainsi l'avis de l'INSEE.

5 Limites du modèle et de sa mise en pratique

Dans cette partie, nous nous intéressons aux éventuelles limites et axes d'amélioration que nous pourrions relever concernant le modèle statistique et sa mise en pratique.

5.1 Limites du modèle statistique

Tout d'abord, nous avons fait l'hypothèse que les individus étaient tous indépendants, et que la mort d'un individu n'avait aucune incidence sur la probabilité de mourir pour un autre individu. De surcroît, nous avons également supposé l'indépendance des différents quotients de mortalité entre les années. Or, cela paraît difficilement envisageable dans un cadre réel. Les "effets de moisson" en sont un bon exemple. Lorsqu'un groupe de personnes subi un nombre important de décès au cours d'une année, elle en connaît généralement

moins l'année d'après (à noter que les auteurs évoque l'effet de moisson dans leur introduction puis n'indique pas l'avoir implémenté dans leur modèle). Cela est notamment le cas pour les personnes âgées et les épidémies de grippe, qui sont plus virulents tous les deux à trois ans. Relever ces hypothèses entraîneraient une forte complexification du modèle, notamment pour les matrices de variance/covariance (les Σ) qui ne seraient alors plus des matrices diagonales, les covariances n'étant alors plus nulles.

Nous avons également supposé que les effectifs de population croissaient au même taux et convergeaient vers un $\lambda_t \in \mathbb{R}$ lorsque nous les faisons tendre vers l'infini, ce qui n'est pas très restrictif.

5.2 Limites de l'application

Tout d'abord, les bases de données que nous avons pu utiliser ne sont pas toujours idéales. En effet, la deuxième base de données sous-estimait très clairement les quotients de mortalité des années précédentes, ce qui ne nous a pas permis d'obtenir de bons résultats. En outre, nous aurions aimé obtenir des données géographiques, pour également étudier la surmortalité en fonction de la zone géographique. Néanmoins, nous n'avons pas trouvé de base de données avec ce genre d'informations.

Nous aurions également aimé avoir d'autres fonctions f pour prédire les quotients de mortalité de 2020. Il est peut-être possible d'implémenter une régression linéaire sur les quotients de mortalité des 10 dernières années pour estimer nos quotients de mortalité de 2020. C'est la nouvelle tendance pour l'estimation de quotient de mortalité et nous aurions aimé avoir le temps d'implémenter ce genre de méthode.

Afin de reproduire ce qui a été fait dans l'article, nous nous sommes restreints à la France métropolitaine et aux personnes âgées de 0 à 99 ans (contrairement à l'INSEE qui considère l'ensemble de la population française). Ainsi, toute comparaison de nos résultats avec ceux de l'INSEE comporte un léger biais. Néanmoins, nous pouvons raisonnablement penser que ces différences altèreraient à la marge les ordres de grandeur de nos résultats.

6 Annexes

6.1 Nos Résultats

Lien du GitHub : https://github.com/JulesChpn/StatApp_Covid.git

Tableau 1a : Evaluation de la surmortalité en 2020 (3 années de référence : 2017-19)

| Classe d'âge | Population en million d'habitants | Décès attendus | Décès observés | Décès en excès sur classe d'âge | Décès en excès par âge |
|--------------|-----------------------------------|-----------------------------|----------------|-----------------------------------|-------------------------------------|
| 0-64 ans | 51,68 (79,36 %) | 93 211 [91 612 ; 94 810] | 92 478 | (0 % [0 % ; 0 %]) | - 705 (-0,76 % [-0,74 % ; -0,77 %]) |
| 65-99 ans | 13,44 (20,64 %) | 536 279 [528 940 ; 543 618] | 561 538 | 17 920 (3,34 % [3,3 % ; 3,39 %]) | 17 013 (3,17 % [3,13 % ; 3,22 %]) |
| Somme | 65,12 (100 %) | 629 490 [628 011 ; 630 970] | 654 016 | 23 046 (3,66 % [3,65 % ; 3,67 %]) | 16 309 (2,59 % [2,58 % ; 2,6 %]) |

Figure 1 : Résultats de l'article

| | Population | Décès attendus | Décès observés | Différence de décès | Pourcentage de différence |
|-----------|------------|------------------|----------------|---------------------|---------------------------|
| (0, 64] | 51821513 | [93716, 100149] | 92478 | [-7671, -1238] | [-7.66 %, -1.32 %] |
| (64, 100] | 13443795 | [484223, 496919] | 561538 | [64619, 77315] | [13.0 %, 15.97 %] |
| Total | 65265308 | [577939, 597068] | 654016 | [56948, 76077] | [9.54 %, 13.16 %] |

Figure 2 : Résultats pour la première base de données et la fonction f_2

| | Population | Décès attendus | Décès observés | Différence de décès | Pourcentage de différence |
|-----------|------------|------------------|----------------|---------------------|---------------------------|
| (0, 64] | 51821513 | [90616, 101739] | 92478 | [-9261, 1862] | [-9.1 %, 2.05 %] |
| (64, 100] | 13443795 | [488035, 509574] | 561538 | [51964, 73503] | [10.2 %, 15.06 %] |
| Total | 65265308 | [578651, 611313] | 654016 | [42703, 75365] | [6.99 %, 13.02 %] |

Figure 3 : Résultats pour la première base de données et la fonction f_1

| | Population | Décès attendus | Décès observés | Différence de décès | Pourcentage de différence |
|-----------|------------|------------------|----------------|---------------------|---------------------------|
| (0, 64] | 51821513 | [90312, 96641] | 92478 | [-4163, 2166] | [-4.31 %, 2.4 %] |
| (64, 100] | 13443795 | [454539, 466901] | 561538 | [94637, 106999] | [20.27 %, 23.54 %] |
| Total | 65265308 | [544851, 563542] | 654016 | [90474, 109165] | [16.05 %, 20.04 %] |

Figure 4 : Résultats pour la deuxième base de données et la fonction f_2

| | Population | Décès attendus | Décès observés | Différence de décès | Pourcentage de différence |
|-----------|------------|------------------|----------------|---------------------|---------------------------|
| (0, 64] | 51821513 | [86210, 97156] | 92478 | [-4678, 6268] | [-4.81 %, 7.27 %] |
| (64, 100] | 13443795 | [443365, 464032] | 561538 | [97506, 118173] | [21.01 %, 26.65 %] |
| Total | 65265308 | [529575, 561188] | 654016 | [92828, 124441] | [16.54 %, 23.5 %] |

Figure 5 : Résultats pour la deuxième base de données et la fonction f_1

6.2 Définitions

Vous trouverez ici quelques définitions importantes pour mieux comprendre l'article de nos auteurs et les thématiques abordées dans ce projet.

Surmortalité : différence entre le nombre de décès observés lors du passage d'un événement sanitaire important et le nombre de décès attendus pour la même période.

Taux de mortalité : pour un groupe et une période donnés, c'est le rapport entre le nombre d'individus du groupe décédés au cours de la période rapporté à la population totale de ce groupe pendant la période.

Nombre attendu de décès : estimé par le nombre de personnes vivantes au début de l'année multiplié par le taux de mortalité attendu. Il se calcule en utilisant un modèle statistique capable de prédire ce nombre attendu.

Comorbidité : association de deux maladies psychiques ou physiques, sans causalité établie, qui dans notre cas, empêche de connaître la véritable raison de la mort d'un patient.

Effet de moisson : lors d'un événement sanitaire fort, les personnes âgées et fragiles meurent de manière ciblée sur une période relativement courte. Cela entraîne un épuisement temporaire de cette catégorie de population, ce qui entraîne une sous-mortalité lors des années suivantes (alternance entre gripes fortement et faiblement meurtrières). Il ne faut donc pas utiliser ces années de « répit » comme référence, ce que l'INSEE fait.

Intervalle de confiance : toute valeur observée comprise dans l'intervalle de confiance de la valeur prédite est considérée comme attendue, quelques soient les différences constatées. En revanche, l'excès ou le défaut de la variable étudiée est la somme des valeurs se situant au-delà des bornes définies par l'intervalle de confiance autour de la valeur prédite.

6.3 Chiffres clés

Vous trouverez ici tous les chiffres clés à avoir en tête pour mieux comprendre les ordres de grandeur et les dynamiques démographiques actuelles:

Augmentation du nombre de décès : Depuis 2004, le nombre de décès augmente en moyenne de 2% par an.

Comorbidité et Covid-19 : 90% des personnes admises en réanimation et 65% des personnes décédées pour motif de Covid-19 présentaient au moins une comorbidité.

Age médian au décès de Covid-19 : Age médian au décès à l'hôpital des malades de l'épidémie de coronavirus est de 85 ans et près de 92,5% des personnes avaient 65 ans et plus.

Population et décès en 2020 : En 2020, les plus de 65 ans représentait 19,45% de la population pour 84,3% des décès. Symétriquement, les moins de 65 ans représentaient 80,5% de la population pour 15,7% des décès.

Taux de mortalité annuel moyen : 1,004% en 2020 ; 0,922% en 2019 ; 0,920% en 2018 ; 0,918% en 2017. En 2020, cela représentait 0,17% pour les moins de 65 ans, 4,18% pour les plus de 65 ans, 35% pour les plus de 95 ans.

Population métropolitaine en 2020 : 65 123 843 individus.

Décès en 2020 : Il y a eu 654 016 décès dont 64 632 décès attribués à la Covid-19.

Surmortalité due au Covid-19 : Le Covid-19 aurait engendré une surmortalité allant de 2,59% à 3,66% en 2020.

6.4 Revue de littérature

Lien 1 : Combien y aurait-il eu de décès en France sans l'épidémie de Covid-19 ?

Natalie Blanpain, INSEE :

Le nombre de décès en France augmente en général chaque année à cause de la croissance de la population et de son vieillissement, mais les risques de mourir à un âge élevé sont en baisse. L'Insee a publié un bilan de l'augmentation des décès en 2020 causée par l'épidémie de Covid-19 (+ 9% par rapport à 2019). L'impact de cette épidémie sur la mortalité peut être évalué en comparant les décès attendus si l'épidémie n'avait pas eu lieu. Cet article propose une méthode afin de calculer ce nombre de décès attendu en 2020.

Les personnes âgées de 65 ans et plus représentent 84% des décès en 2019, alors que leur part dans la population est 4 fois plus faible. Cependant, les taux de mortalité ont généralement diminué au fil des ans en raison de progrès dans divers domaines, ce qui signifie que les gens vivent plus longtemps. L'épidémie de Covid-19 en 2020 a interrompu cette tendance à la baisse des quotients de mortalité.

Le vieillissement de la population tend à augmenter le nombre de décès, tandis que la baisse des quotients de mortalité tend à les réduire. Entre 2010 et 2019, la première tendance a dominé la seconde, entraînant une augmentation des décès. Les méthodes pour prédire les décès en 2020 en utilisant les quotients de mortalité antérieurs seuls peuvent sous-estimer les décès attendus en raison de la tendance à la baisse des quotients de mortalité et de l'augmentation de l'espérance de vie.

L'objectif du texte est d'estimer le nombre de décès qui auraient eu lieu en 2020 sans l'épidémie de Covid-19. Pour ce faire, les démographes supposent que les quotients de mortalité continueraient d'évoluer à la même allure qu'au cours des dernières années. En se basant sur la période 2010-2019 (+0,8 mois d'espérance de vie par an), on aurait attendu 621 900 décès, soit 8 700 de plus qu'en 2019. Cette hausse vient du vieillissement de la population (+13 800 décès) et du jour en plus de l'année bissextile 2020 (+1 900 décès) nuancée par la baisse des quotients de mortalité (-7 000 décès). Cependant, en 2020, 669 000 décès ont finalement été comptabilisés, soit 47 100 de plus que prévu.

Cependant, le nombre de décès attribué à la COVID-19 est supérieur à l'impact de la crise sanitaire sur les décès totaux, car il inclut les décès de personnes vulnérables qui

seraient décédées en 2020 même sans l'épidémie, mais qui ont été anticipés en raison de l'épidémie. L'impact de la crise sanitaire, qui mesure l'écart entre le nombre de décès attendus et observés, ne prend pas en compte ces décès anticipés, mais il est réduit grâce à des décès évités. Les statistiques officielles définitives des causes médicales de décès sont nécessaires pour préciser la comparaison entre le nombre de décès attribués à la COVID-19 et l'impact de la crise sanitaire sur les décès totaux.

Lien 2 : 2020 : une hausse des décès inédite depuis 70 ans

Sylvie Le Minez, Valérie Roux, INSEE :

Cet article rapporte que les deux vagues successives de la Covid-19 en 2020 n'ont pas eu la même durée ni la même intensité sur les différentes régions en France métropolitaine et d'outre-mer. La première vague est survenue au printemps 2020, où il y a eu une hausse de 27% des décès (+27 300 morts) sur la période mars/avril par rapport à 2019. Le pic de décès durant cette période a été atteint le 1er avril 2020 avec 2810 décès en un jour, puis il a diminué sur le mois d'avril. Entre mai et août 2020, le nombre de décès se situe à nouveau dans la moyenne des années précédentes. La deuxième vague qui a duré de septembre à décembre 2020 a été moins intense, mais a duré plus longtemps (+34 300 morts), soit une augmentation de 17% par rapport à la même période en 2019. Le pic de mort est atteint le 7 novembre avec 2340 décès en un jour.

De plus, les deux vagues successives ont eu impact différent suivant les régions. En effet, la première vague a été plus meurtrière en Île-de-France et dans le Grand Est que partout en France. La première vague a commencé dans les Hauts-de-France en mars 2020 avec un pic de +61% dans l'Oise. Dans le Grand Est, entre mars et avril, les décès ont été supérieurs de 55% à ceux de 2019. En Bourgogne Franche-Comté, l'augmentation de la mortalité a été deux fois moindre que dans le Grand Est au printemps et équivalente à celle des Hauts-de-France. Enfin, après des débuts timides, à partir du 16 mars, l'Île-de-France a enregistré une hausse de la mortalité beaucoup plus importante que partout ailleurs, avec une surmortalité de 91% entre mi-mars et avril.

Pendant l'automne 2020, la deuxième vague de Covid-19 a été plus forte dans les régions qui avaient été moins touchées au printemps, en particulier Auvergne-Rhône-Alpes (+38%), où l'excédent est moins important que celui de l'Île-de-France mais a duré plus longtemps. Les autres régions du centre et du Sud de la France, comme la Bourgogne-Franche-Comté et la région Provence-Alpes-Côte d'Azur ont également été particulièrement touchés par cette seconde vague.

Dans les DOM, les excédents de mortalité varient énormément suivant les départements. A Mayotte par exemple, cet excédent de morts est de 25% sur toute l'année 2020, par rapport à 2019, alors qu'il est nul pour la Martinique et la Réunion, et il est même négatif pour la Guyane.

Pour conclure, l'article stipule que en 2020, le nombre de décès a augmenté dans la plupart des pays européens. La France se situe dans une position médiane avec une hausse

de 9% des décès. Les pays limitrophes de la France ont enregistré une hausse plus importante des décès, sauf l'Allemagne et le Luxembourg. Sept pays européens ont enregistré une hausse supérieure ou égale à 14% des décès : l'Espagne, la Pologne, la Belgique, la Slovaquie, la Bulgarie, la République tchèque et l'Italie. Au printemps, lors de la première vague de la pandémie de Covid-19, la mortalité n'a augmenté que dans sept pays en Europe, mais avec la seconde vague à l'automne, la crise sanitaire s'est répandue plus largement.

References

- [1] Laurent Toubiana, Laurent Mucchielli, Pierre Chaillot, Jacques Bouaud (2022). L'épidémie de Covid-19 a eu un impact relativement faible sur la mortalité en France : <https://docplayer.fr/216724313-L-epidemie-de-covid-19-a-eu-un-impact-relativement-faible-sur-la-mortalite-en-france.html?fbclid=IwAR2L7bhpnR2USVWq7QXPuHRi1iiisuweZcoUM5o7sEoeQNRw99ls73dNog>
- [2] Nathalie Blanpain, INSEE (2022). Combien y aurait-il eu de décès en France sans l'épidémie de Covid-19 ? <https://blog.insee.fr/combien-y-aurait-il-eu-de-deces-en-france-sans-lepidemie-de-covid-19/>
- [3] Sylvie Le Minez, Valérie Roux, INSEE (2022). 2020 : une hausse des décès inédite depuis 70 ans : <https://www.insee.fr/fr/statistiques/5347349?sommaire=5413896>
- [4] OMS (2021). COVID-19 – Chronologie de l'action de l'OMS : <https://www.who.int/fr/news/item/27-04-2020-who-timeline---covid-19>
- [5] AFP (2012). La grippe sature les urgences : https://www.lexpress.fr/actualite/societe/sante/la-grippe-sature-les-urgences_1086604.html