

Life Tables

2025-01-29

Devoir II : Etude de tables de mortalite

! Attendu

Un fichier `nom1_nom2.Rmd` ou `nom1_nom2.qmd` au format Rmarkdown (`.Rmd`) ou Quarto (`.qmd`), compilable (mais pas compilé) en format `html`, à charger sur [Moodle](#).

Dans ce fichier `nom1_nom2.xxx`, on trouvera le code nécessaire à la génération des graphiques et des tables correspondants aux questions ci-dessous.

Date de rendu : 29 mai 2024 à 23h59

À réaliser en binôme.

Tables de mortalité (1900-1925)

Ce devoir porte sur les tables de mortalité (*life tables*) américaines et européennes entre 1900 et 1925.

Les tables ont été obtenues de <https://www.mortality.org>.

Nous étudions les tables de mortalité de quelques pays d'Europe occidentale (France, Grande-Bretagne –en fait, Angleterre et Pays de Galles–, Italie, Pays-Bas, Espagne, et Suède) et des États-Unis d'Amérique.

Les tables peuvent être téléchargées à l'aide des instructions suivantes :

```
datafile <- 'full_life_table.Rds'
fpath <- stringr::str_c("./DATA/", datafile)

# here::here('DATA', datafile)
# check getwd() if problem
```

```

if (! file.exists(fpath)) {
  download.file("https://stephane-v-boucheron.fr/data/full_life_table.Rds",
               fpath,
               mode="wb")
}

life_table <- readr::read_rds(fpath)

life_table <- life_table %>%
  mutate(Country = as_factor(Country)) %>%
  mutate(Country = fct_relevel(Country, "Spain", "Italy", "France",
                              "England & Wales", "Netherlands", "Sweden", "USA")) %>%
  mutate(Gender = as_factor(Gender))

life_table <- life_table %>%
  mutate(Area = fct_collapse(Country,
                             SE = c("Spain", "Italy", "France"),
                             NE = c("England & Wales", "Netherlands", "Sweden"),
                             USA="USA"))

```

Sur les données françaises, le document [Tables de mortalité françaises pour les XIXe et XXe siècles et projections pour le XXIe siècle](#) contient des informations utiles sur la construction des tables de mortalité.

Notation (Rappel)

Dans la suite F désigne une fonction de répartition sur $\mathbb{N} = \mathbb{Z}_+$, et $\bar{F} = 1 - F$ la fonction de survie associée. Cette fonction F est définie à partir des quotients de mortalité (voir plus bas). Elle ne décrit pas la pyramide des âges. On définit une fonction pour chaque année t . Pour chaque année, pays, sexe, $F_t(x)$ est la proportion des membres d'une cohorte (fictive) qui vivent au moins jusqu'à l'âge x dans l'année t .

qx (age-specific) risque de décès à l'âge (révolu) x , ou encore *quotient de mortalité* à l'âge x pour l'année t : $q_{t,x} = \frac{\bar{F}_t(x) - \bar{F}_t(x+1)}{\bar{F}_t(x)}$.
 Pour chaque année, chaque âge, $q_{t,x}$ est déterminé par les données de l'année.

Nous avons aussi

$$\bar{F}_t(x+1) = \bar{F}_t(x) \times (1 - q_{t,x+1}).$$

mx *taux central de décès* à l'âge (révolu) x durant l'année t . C'est relié à $q_{t,x}$ par

$$m_{t,x} = -\log(1 - q_{t,x}),$$

ou de manière équivalente $q_{t,x} = 1 - \exp(-m_{t,x})$.

lx la *fonction de survie*: un multiple de proportion de personnes encore vivantes à l'âge x . Ces valeurs sont calculées à partir de $q_{t,x}$ via la formule

$$l_t(x+1) = l_t(x) \times (1 - q_{t,x}),$$

avec $l_{t,0}$, la racine (*radix*) de la table, en fait, choisi égal à 100000. Les fonctions $l_{t,\cdot}$ et \bar{F}_t sont liées par

$$l_{t,x+1} = l_{t,0} \times \bar{F}_t(x).$$

dx $d_{t,x} = q_{t,x} \times l_{t,x}$

Tx Nombre total de personnes-années vécues par la cohorte des gens d'âge compris entre x et $x+1$ (pour une année donnée dans une société donnée). C'est nombre d'années vécues par les $l_{t,x+1}$ personnes qui survivent à l'intervalle de temps, et les $d_{t,x}$ personnes qui décèdent durant cette intervalle (ici l'intervalle est une année). Les premiers contribuent chacun exactement 1 année, alors que ces derniers contribuent, en moyenne, approximativement pour une demi-année. Ainsi $L_{t,x} = l_{t,x+1} + 0.5 \times d_{t,x}$. Cette approximation équivaut à supposer qu'un décès à l'âge révolu x , intervient en moyenne au milieu de l'année. C'est acceptable excepté durant la première année (âge 0) et aux grands âges. Nous en restons à l'approximation simpliste $L_{t,x} = l_{t,x+1}$.

ex: Espérance de vie *résiduelle* à l'âge x pour l'année t . C'est (presque) l'espérance de la loi sur $[0, \infty)$ définie par F_t (et donc par les **qx**), de la façon suivante: si $X \sim F$, c'est

$$\mathbb{E}_{\{X \geq x\}} [X - x] = \frac{\mathbb{E} [(X - x) \mathbb{I}_{X \geq x}]}{\mathbb{E} [\mathbb{I}_{X \geq x}]}$$

Tip

Le package R nommé **demography** met à disposition un certain nombre d'outils et de concepts élaborés par les démographes.

Sources: *Demography: measuring and modeling population processes*. Preston, Heuveline et Guillot. Blackwell Publishing. 2001.

Question

Pour chaque pays et chaque sexe, illustrer et commenter (brièvement) l'évolution des quotients de mortalité entre 1900 et 1913.

Remarquer qu'on peut étudier **qx** comme une fonction de l'année t , mais aussi pour une année donnée, étudier **qx** comme une fonction de l'âge x .





Question

Pour chaque pays, chaque sexe, chaque année entre 1900 et 1913, puis entre 1921 et 1925, effectuer une *régression linéaire* du logarithme du quotient de mortalité en fonction de l'âge, pour les âges compris entre 30 et 70 ans. Illustrer et commenter.

Question

Pour chaque pays et chaque sexe, considérer la *cohorte* des individus nés en 1890. Déterminer les quotients de mortalité effectivement subis par cette cohorte entre 1890 et 1980. Illustrer la différence entre les quotients de mortalité tirés des *tables du moment* de l'année 1890 et les quotients de mortalités effectivement subis.

Barème

Critère	Points	Détails
Orthographe et grammaire	20%	English/Français 
Graphiques	25%	Choix des aesthetics , geom , scale ... 
Style des Graphiques	15%	Titres, légendes, étiquettes ... 
Manipulations de tables	25%	
Respect DRY	15%	Principe DRY < fab wikipedia-w > Wikipedia

Caution

Ceci n'est pas un devoir d'Histoire. Ne cherchez pas à montrer votre culture. Demandez-vous ce qu'il y a de remarquable dans les données, et énoncez les questions que ces données peuvent poser aux historiens.