

# Projet Actuariat Vie



Amira R'houma

Nada Itaief

Ala Bayoudh

Wajd Meskini

**Mohamed Hammami** 

# Sommaire

Introduction	3
Chapitre I: Cas d'étude: La Norvège	4
1.1 Motivation:	4
1.2 Mortalité en fonction de l'âge et en fonction des années	5
1.2.1 Mortalité sur l'ensemble des années	5
1.2.2 Mortalité sur un ensemble prédéfini	7
1.2.3 Etude comparative: La France	7
1.2.4 Etude comparative: Les sexes	8
1.2.5 Le log des taux de mortalité en 3D.	10
1.3 Taux de mortalité : Etude des cohortes	11
1.3.1 Cohortes sur l'ensemble des années	12
1.3.2 Cohortes sur les années entre 1895 et 1905	13
1.3.3 Cohortes sur les années entre 1990 et 2010	14
1.4 Lissage	15
1.4.1 Spline monotone	15
1.4.2 Spline Standard	16
1.4.3 Spline Concave	17
1.4.4 Lissage localement quadratique	18
1.4.5 Etude comparative	19
Chapitre II: Modèle de Lee Carter pour les données de mortalité	20
2.1 Hypothèse centrale	20
2.2 Cas d'étude: Le Royaume-Uni	21
2.2.1 Application du modèle de Lee Carter	22
2.2.2 Paramètre ax	
2.2.3 Paramètre <i>bx</i>	22
2.2.4 Etude l'Ecart absolu des coefficients	23
2.2.5 Paramètre kt	24
2.3 Extrapolation des $(kt)t$	25
Conclusion générale	25
Bibliographie	25

Table des figures	
Figure 1: Norvège Mortalité Totale	6
Figure 2: Norvège MT Années prédéfinies	7
Figure 3: France MT années prédéfinies	
Figure 4: Norvège MT mortalité féminine	9
Figure 5: Norvège MT mortalité masculine	10
Figure 6: Norvège MT 3D	11
Figure 7: Norvège : Cohortes	13
Figure 8 : Norvège Cohortes 1895 1905	
Figure 9: Norvège Cohortes 1990 2010	
Figure 10 : Norvège MT 2003 Spline Monotone	16
Figure 11: Norvège MT 2003 Spline Standard	17
Figure 12: Norvège MT 2003 Spline Concave	18
Figure 13: Norvège MT 2003 lissage localement quadratique	19
Figure 14: Norvège MT 2003 Analyse comparative	20
Figure 15: Royaume-Uni Mortalité totale	21
Figure 16: Coef.ax sur données britanniques	22
Figure 17: Coef bx sur données britanniques	23
Figure 18: Ecart avec population totale	24
Figure 19: Coef kt sur données britanniques	25

### Introduction

Dans les domaines de la science actuarielle, de la santé, de l'épidémiologie et de la planification nationale, on ne peut nier l'importance des données de mortalité. Les niveaux de mortalité reflètent la qualité de vie en quantité. La prévision de la population est essentielle pour toute planification à long terme de la prestation de services d'une nation. Par conséquent, l'élaboration d'un modèle de prévision du taux de mortalité peut aider un pays à développer sa qualité de vie.

L'un des "privilèges nécessaires" de l'expérience "vie" chez l'humanité est de vivre une vie longue et saine. Les progrès récents dans la technologie et dans la médecine ont fourni un soutien énorme pour satisfaire cette nécesté. Cependant, de grandes disparités sont visibles dans les niveaux de mortalité entre les pays et les régions. La réduction de la mortalité, en particulier de la mortalité infantile et maternelle, fait partie des objectifs de développement convenus au niveau international au XXIe siècle .

Avec l'augmentation de l'espérance de vie à la naissance, il devient plus que jamais intéressant de trouver des modèles quantitatifs des taux de mortalité. Une fois que nous sommes en mesure de modéliser le vieillissement humain, nous sommes en mesure de mieux évaluer l'influence des progrès de la médecine sur la mortalité et donc de mieux maîtriser la condition humaine.

Cette étude rentre dans un cadre académique, et est axée sur la visualisation et modélisation des taux de mortalité en utilisant le modèle Lee Carter pour le Royaume-Uni et la Norvège, pays ayant un niveau de vie parmi les plus élevés sur la base de l'indice de développement humain. Le modèle prédit peut être utilisé pour construire des tables de mortalité pour les pays concernés et peut également être utilisé dans des applications scientifiques actuarielles et dans la planification de régimes de retraite.

la méthodologie et le cadre statistique de l'analyse sont discutés. Nous décrivons le modèle de Lee Carter pour la modélisation et la prévision des taux de mortalité. Les théories utilisées sont également discutées dans cette section. Dans la section des résultats et des discussions, la méthodologie est illustrée par l'analyse des données sur la mortalité en Norvège.

Les graphiques linéaires sont utilisés pour visualiser le profil de mortalité de chaque groupe d'âge et modèle de naissance de la Norvège. À travers les graphiques, les valeurs aberrantes peuvent être observées clairement. Enfin. Le logiciel statistique R a été utilisé pour tous les calculs de cet article.

# Chapitre I: Cas d'étude: La Norvège

## 1.1 Motivation:

Depuis la déclaration de neutralité de la Suède (dont la Norvège a fait partie durant un sous la forme des Royaumes unis de Suède et de Norvège), la Norvège a connu une stabilité économique et politique exemplaire, et même l'interlude d'occupation lors de la Deuxième

Guerre Mondiale n'y a pas changé grand-chose puisque la Norvège était loin des principales lignes de front et ne constituait pas un objectif stratégique important. De fait, la Norvège a été un état stable et fonctionnel pendant toute la durée d'observation, et un état stable donne des données fiables.

Chargeons les taux bruts de mortalité (téléchargés du site HMD <a href="http://www.mortality.org/">http://www.mortality.org/</a>). La fonction read.demogdata() lit les taux de mortalité et les effectifs sous risque.

Ces deux quantités sont liées par l'égalité suivante:

$$\mu_{x,t} = \frac{d_{x,t}}{E_{x,t}}$$

Avec comme paramètres:

 $d_{x,t}$ : Nombre de décès à l'âge x et à l'année calendaire t

 $E_{x,t}$ : Nombre total de vies vécues entre les âges x et x+1

L'objet contient une liste des années de la période d'observation (dans notre cas, 1846 et 2014), une liste contenant les âges possibles (0 à 110), une matrice contenant la table des taux de mortalité pour les hommes, les femmes et une table combinée, et une liste contenant la population pendant chaque année de la période d'observation.

# 1.2 Mortalité en fonction de l'âge et en fonction des années

On s'intéresse à visualiser le log du taux de mortalité en fonction de l'âge et en fonction des années.

#### 1.2.1 Mortalité sur l'ensemble des années

```
max.age=max(thor$age)
,plot.type="functions", main="Norvège: Mortalité totale")
```

# Norvège: Mortalité totale

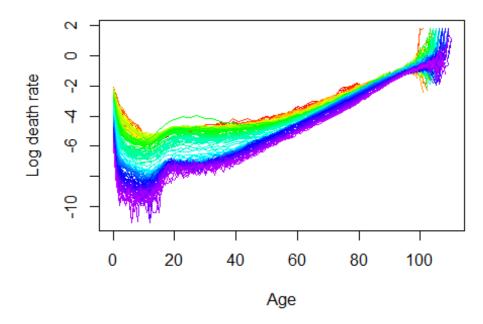


Figure 1: Norvège Mortalité Totale

Le résultat n'est pas très lisible, mais on remarque tout de même une abondance des bleus et violets (années élevées, deuxième moitié du 20ème siècle, début 21ème siècle) ayant des taux de mortalités faibles et une abondance de rouges et verts (années faibles, fin 19ème siècle, première moitié du 20ème siècle).

Pour y remédier, on va réduire le nombre d'années à quatres périodes de l'histoire - qui n'ont pas grand effet sur la Norvège.

La première concerne les générations nées à temps pour pouvoir participer ou vivre la Première Guerre mondiale (1895-1905).

La seconde concerne les générations nées à temps pour pouvoir participer ou vivre la Deuxième Guerre mondiale (1915:1925).

La troisième concerne les premières naissances de l'après-guerre, nées à temps pour assister à la crise pétrolière de 1971, et les guerres de décolonisation (1945-1960). La Norvège n'ayant aucune colonie et possédant des réserves généreuses de pétrole, elle a été épargnée tous ces problèmes.

La quatrième concerne les premières générations à vivre dans un monde sans guerre froide (1980-1990).

## 1.2.2 Mortalité sur un ensemble prédéfini

# Norvège: Mortalité totale

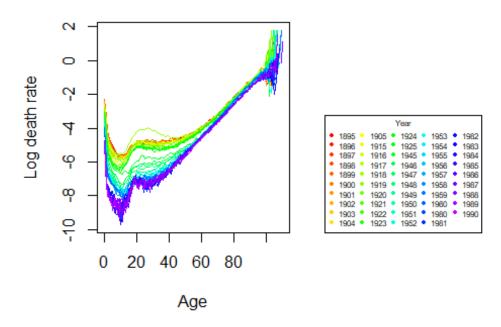


Figure 2: Norvège MT Années prédéfinies

On confirme les interprétations précédentes, avec une légère augmentation pour les générations ayant assisté à la Deuxième Guerre Mondiale et l'Occupation. Il serait encore plus judicieux de faire une analyse comparative avec la France, qui a vécu tous les conflits concernés.

## 1.2.3 Etude comparative: La France

## France: Mortalité totale

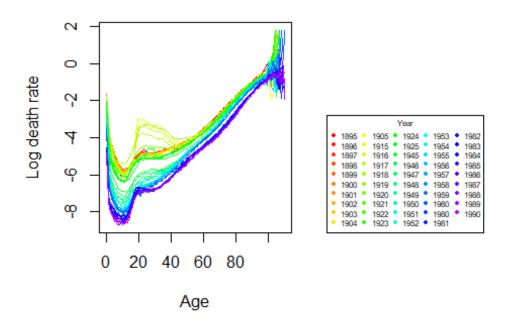


Figure 3: France MT années prédéfinies

La différence est flagrante, surtout pour les générations les plus jeunes à avoir eu l'infortune d'être nés à temps pour la Première Guerre Mondiale, avec des taux de mortalités très élevés pour des hommes et des femmes fraîchement sortis de l'adolescence vers les tranchées.

Il serait également approprié de voir le comportement de la mortalité selon les sexes, en utilisant les mêmes dates comme repaires par souci de lisibilité.

# Norvège: Mortalité féminine

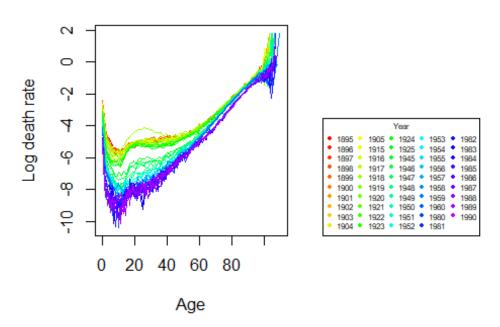


Figure 4: Norvège MT mortalité féminine

# Norvège: Mortalité masculine

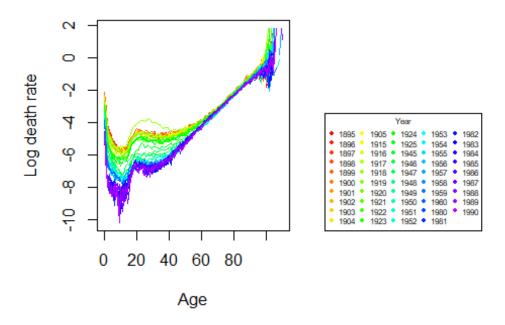


Figure 5: Norvège MT mortalité masculine

Il est intéressant de mentionner que le taux de mortalité féminine présente généralement une courbe en cloche inversée: Décroissante jusqu'à un minimum au niveau des 10 ans, et croissante ensuite. La courbe de mortalité masculine présente une courbe en double cloche, avec un minimum, un maximum local, un minimum local et une croissance.

La courbe de mortalité féminine comporte également une volatilité importante. Les femmes sont également épargnées du pic de mortalité que la courbe rencontre chez les hommes à la sortie de l'adolescence. Cela est probablement dû au fait que les normes culturelles imposent souvent aux hommes une tendance à prendre plus de risques pour s'affirmer à la sortie de l'adolescence, en une sorte de rite de passage.

## 1.2.5 Mortalité en fonction de l'âge et en fonction des années en 3D.

On s'intéresse à la visualisation du log des taux de mortalité en fonction de l'âge et en fonction des années en 3D.

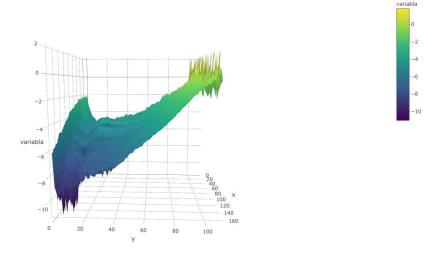


Figure 6: Norvège MT 3D

La visualisation en 3D ne rajoute rien de particulier à nos interprétations précédentes, à part le fait qu'elle représente mieux la différence entre les mortalités des différentes années de naissance.

## 1.3 Taux de mortalité : Etude des cohortes

Les données précédentes  $\mu_{x,t}$  ont été affichées en fonction de l'âge  $(x,\mu_{x,t})$  et en fonction de l'année  $(t,\mu_{x,t})$  en considérant la matrice  $(\mu_{x,t})_{x,t}$ . Cependant dans cette lecture longitudinale, on mélange les générations, autrement dit, les cohortes, dans ces graphiques puisque les individus nés en 1901 (par exemple) ont 1 an en 1902, 2 an en 1903,... Mais on ne les affiche jamais sur un même graphique.

En assurance vie, le principe est de suivre un individu par année de naissance pour pouvoir offir un produit d'assurance-vie pour les individus de la même génération. Ces trois dimensions x (âge), t (date) et t-x (année de naissance) n'est pas sans rappeler la lecture des triangles de provisionnement j (développement, ou âge d'un sinistre), i+j (année calendaire, ou date de paiement) et i (année de survenance, ou année de naissance du sinistre).

La génération est obtenue par g=x-t, par exemple 1901=1902-1=1903-2=...

Pour extraire une génération, nous avons récupéré la fonction suivante.

```
cohort <- function(year, rates, log=FALSE)
{
    xg <- diag(rates[, colnames(rates) >= year])
    names(xg) <- paste(year, rownames(rates)[1:length(xg)], sep="-x=")
    if(log)
        xg <- log(xg)
    xg
}</pre>
```

Evidemment, nous ne nous intéresserons pas à une seule cohorte, plutôt à une cohorte de cohortes. Nous avons donc pris la liberté de rajouter une fonction pour accomplir précisément cela.

```
cohort_of_cohorts <- function(years, rates, log=FALSE)
{
    xgs<-vector()
    for (year in years)
    {
        xg <- diag(rates[, colnames(rates) >= year])
        names(xg) <- paste(year, rownames(rates)[1:length(xg)], sep="-x=")
    if(log)
        xg <- log(xg)
        xgs<-c(xgs,list(xg))
    }
    xgs
}</pre>
```

Nous nous intéresserons donc à la visualisation de ces cohortes sur un graphe similaire aux graphes précédents.

## 1.3.1 Cohortes sur l'ensemble des années

```
colfunc <- colorRampPalette(c("red", "blue"))
cohorts<-cohort_of_cohorts(thor$year, thor$rate$total, log=T)

plot(cohort(thor$year[1], thor$rate$total,
log=T),col=colfunc(length(thor$year))[1], type="l", ylim=c(-11,5),
main="Norvège: Cohortes")
i<-2
for (one_cohort in cohorts){
   lines (one_cohort, col=colfunc(length(thor$year))[i])
   i<-i+1
}</pre>
```

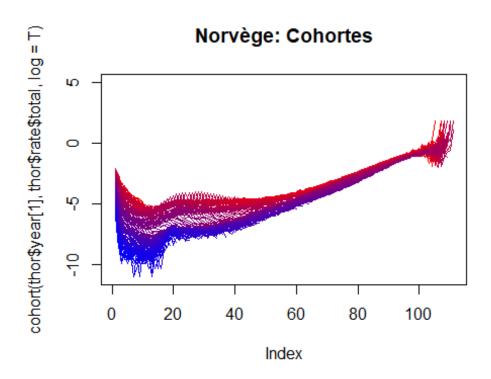


Figure 7: Norvège : Cohortes

Par souci de netteté, nous n'allons visualiser que les années entre 1895 et 1905.

### 1.3.2 Cohortes sur les années entre 1895 et 1905

```
colfunc <- colorRampPalette(c("red", "blue"))
cohorts<-cohort_of_cohorts(1896:1905, thor$rate$total, log=T)

plot(cohort(1895, thor$rate$total, log=T),col=colfunc(10)[1],
type="l",ylim=c(-6,4), main="Norvège: Cohortes 1895 1905")
i<-2
for (one_cohort in cohorts){
    lines (one_cohort, col=colfunc(10)[i])
    i<-i+1
}
legend(legend=unique(1895:1905),
        col=colfunc(10), ncol=5, pch=19,
        title="Year", cex=0.5, "bottomright")</pre>
```

# Norvège: Cohortes 1895 1905

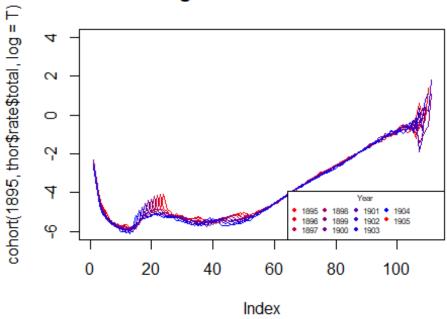


Figure 8 : Norvège Cohortes 1895 1905

## 1.3.3 Cohortes sur les années entre 1990 et 2010

```
colfunc <- colorRampPalette(c("red", "blue"))
cohorts<-cohort_of_cohorts(1991:2010, thor$rate$total, log=T)

plot(cohort(1990, thor$rate$total, log=T),col=colfunc(20)[1],
type="l",ylim=c(-11,-4), main="Norvège: Cohortes 1990 2010")
i<-2
for (one_cohort in cohorts){
   lines (one_cohort, col=colfunc(20)[i])
   i<-i+1
}
legend(legend=unique(1990:2010),
        col=colfunc(20), ncol=5, pch=19,
        title="Year", cex=0.5, "topright")</pre>
```

# Norvège: Cohortes 1990 2010

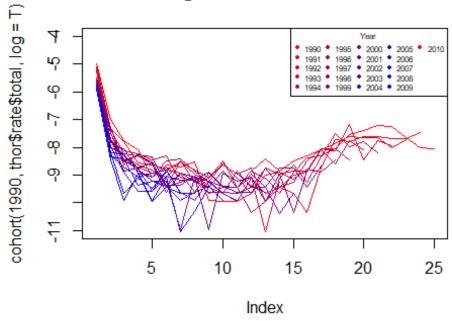


Figure 9: Norvège Cohortes 1990 2010

Par construction dans cette lecture transversale, tous les individus n'auront pas été observés pendant la même durée. Les personnes nées en 1990 n'ont que 28 ans. Cela nous mène nécessairement à la nécessité de prédire la mortalité des générations les plus jeunes.

# 1.4 Lissage

On récupère pour des raisons de temps de calcul les années allant de la fin de la Première Guerre Mondiale à la fin de la série.

```
thor20 <- extract.years(thor, years=1918:2014)
```

## 1.4.1 Spline monotone

on commence par le spline monotone.

```
thor.smth.m <- smooth.demogdata(thor20, method="mspline")
plot(thor,series=names(thor$rate)[3], years=2003, type="p", pch=21, ylim=c(-
11, 1), main="Norvège: MT 2003 - Spline Monotone")
lines(thor.smth.m, years=2003, lty=1, col="blue")
legend("topleft",col=c("blue") ,lty=1, leg=c("Spline monotone"))</pre>
```

# Norvège: MT 2003 - Spline Monotone

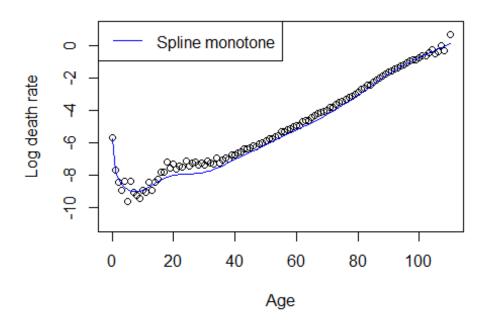


Figure 10 : Norvège MT 2003 Spline Monotone

## 1.4.2 Spline Standard

```
thor.smth.u <- smooth.demogdata(thor20, method="spline")
plot(thor, series=names(thor$rate)[3],years=2003, type="p", pch=21, ylim=c(-
11, 1), main="Norvège: MT 2003 - Spline Standard")
lines(thor.smth.u, years=2003, lty=2, col="red")
legend("topleft",col=c("red") ,lty=2, leg=c("Spline standard"))</pre>
```

# Norvège: MT 2003 - Spline Standard

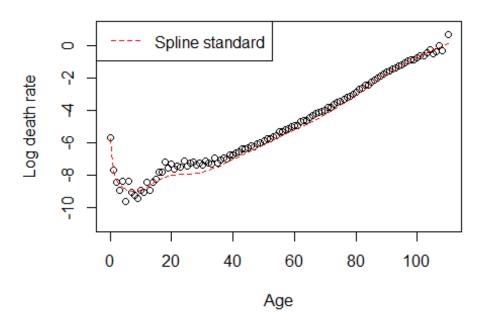


Figure 11: Norvège MT 2003 Spline Standard

## 1.4.3 Spline Concave

```
thor.smth.c <- smooth.demogdata(thor20, method="cspline")
plot(thor, series=names(thor$rate)[3],years=2003, type="p", pch=21, ylim=c(-
11, 1), main="Norvège: MT 2003 - Spline concave")
lines(thor.smth.c, years=2003, lty=3, col="green")
legend("topleft",col=c("green"),lty=2, leg=c("Spline concave"))</pre>
```

# Norvège: MT 2003 - Spline concave

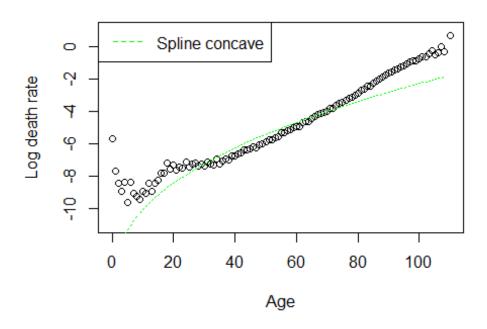


Figure 12: Norvège MT 2003 Spline Concave

## 1.4.4 Lissage localement quadratique

```
thor.smth.q <- smooth.demogdata(thor20, method="loess")
plot(thor,series=names(thor$rate)[3], years=2003, type="p", pch=21, ylim=c(-
11, 1), main="Norvège: MT 2003 - Lissage localement quadratique")
lines(thor.smth.q, years=2003, lty=4, col="black")
legend("topleft",col=c("black"),lty=2, leg=c("Lissage localement
quadratique"))</pre>
```

# Norvège: MT 2003 - Lissage localement quadratiqu

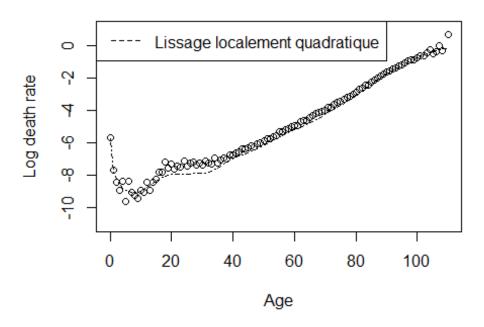


Figure 13: Norvège MT 2003 lissage localement quadratique

## 1.4.5 Etude comparative

```
plot(thor, years=2003, type="p", pch=21, ylim=c(-12, 1), main="Norvège: MT
2003 - Lissage")
lines(thor.smth.m, years=2003, lty=1, col="blue")
lines(thor.smth.u, years=2003, lty=2, col="red")
lines(thor.smth.c, years=2003, lty=3, col="green")
lines(thor.smth.q, years=2003, lty=4, col="black")
legend("topleft",col=c("blue","red","green","black") ,lty=1:4,
leg=c("mspline", "spline", "cspline", "loess"))
```

# Norvège: MT 2003 - Lissage

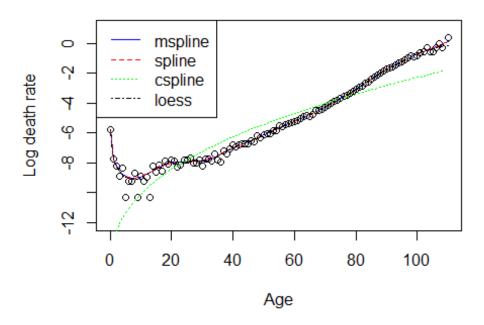


Figure 14: Norvège MT 2003 Analyse comparative

On remarque ici que le lissage par spline localement quadratique représente le mieux la variation du taux de mortalité, suivi par les splines monotones et standard. Le spline concave n'a pas donné de résultat concluant.

# Chapitre II: Modèle de Lee Carter pour les données de mortalité

# 2.1 Hypothèse centrale

Le modèle Lee-Carter est une formule numérique utilisée dans la prévision de la mortalité et la prévision de l'espérance de vie. L'entrée dans le modèle est une matrice de taux de mortalité spécifiques à l'âge, ordonnée de façon monotone par le temps, généralement avec des âges en colonnes et des années en rangées. La sortie est une autre matrice prévue des taux de mortalité . Le modèle de Lee Carter pour les données de mortalité est le suivant :

L'hypothèse centrale du modèle de Lee-Carter est

$$log(\mu_{x,t}) = a_x + b_x k_t + \epsilon_{x,t},$$

pour  $t = 1, \ldots, n$  et  $x = x_{min}, \ldots, x_{max}$ 

Avec comme paramètres

 $a_x$ : Impact de l'âge sur les log-taux au cours du temps. Rend compte du comportement moyen de la mortalité au cours du temps

 $b_x$ : Composante traduisant la sensibilité de la mortalité instantanée par rapport à l'évolution temporelle  $k_t$ . Représente l'évolution croisée avec le temps du taux de mortalité

 $k_t$ : Traduit l'évolution temporelle du taux de mortalité

 $\epsilon_{x,t}$ : Un résidu aléatoire i.i.d. de loi  $N(0,\sigma^2)$ 

L'idée du modèle est d'ajuster à la série des logarithmes des taux instantanés de mortalité (doublement indexée par x et t) une structure paramétrique déterministe à laquelle s'ajoute une suite de variables aléatoires i.i.d. suivant la loi normale centrée et de variance commune. Le critère d'optimisation retenu va consister à maximiser la variance expliquée par le modèle, ce qui revient à minimiser la variance des erreurs.

# 2.2 Cas d'étude: Le Royaume-Uni

# Royaume-Uni: Mortalité totale

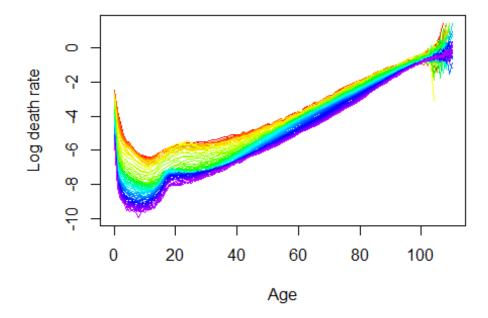


Figure 15: Royaume-Uni Mortalité totale

Nous passons à l'application du modèle de Lee Carter sur notre jeu de données.

## 2.2.1 Application du modèle de Lee Carter

```
lca.male <- lca(arthur, series="male", max.age=100)
lca.female <- lca(arthur, series="female", max.age=100)
lca.total <- lca(arthur, series="total", max.age=100)</pre>
```

Nous avons appliqué le modèle séparément sur les jeux de données totaux, hommes et femmes.

## **2.2.2** Paramètre $a_x$

```
plot(lca.female$ax, main="Coef. ax sur données britanniques", xlab="Age",
ylab="ax", type="l")
lines(x=lca.male$age, y=lca.male$ax, main="ax", lty=2)
legend("bottomright", c("Female", "Male"), cex=0.8, lty=1:2)
```

## Coef. ax sur données britanniques

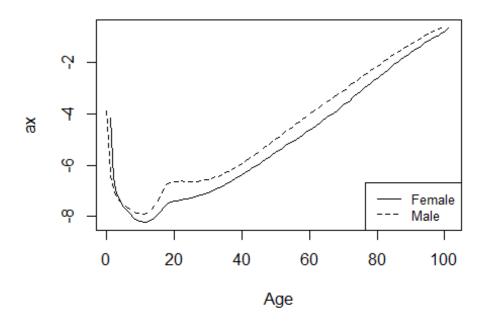


Figure 16: Coef.ax sur données britanniques

Les paramètres  $a_x$  représentent la tendance liée à l'effet isolé de l'âge sur les taux de mortalité. La courbe des  $a_x$  suit la tendance des courbes des données empiriques. Les âges faibles ont une décroissance jusqu'à atteindre un minimum absolu atteint au niveau de l'âge de dix ans, une croissance puis une forte croissance jusqu'à la fin de la vie.

```
2.2.3 Paramètre b_x
```

```
plot(lca.female$bx, main="Coef. bx sur données britanniques",
ylim=c(0,0.03),xlab="Age", ylab="bx", type="l")
```

```
lines(x=lca.male$age, y=lca.male$bx, main="ax", lty=2)
legend("topright",c("Female","Male"), cex=0.8, lty=1:2)
```

## Coef. bx sur données britanniques

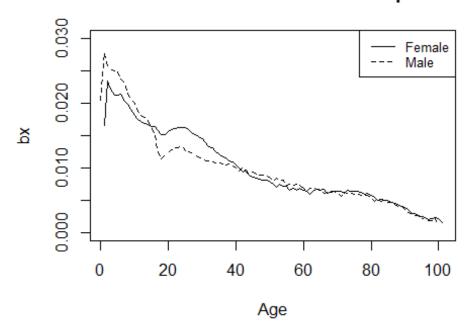


Figure 17: Coef bx sur données britanniques

Les paramètres  $b_x$  représentent l'interaction de l'effet des années calendaires sur les taux de mortalité. Cet effet est toujours positif mais la valeur ne cesse de diminuer avec l'âge. Autrement dit, l'effet des années calendaires agit majoritairement avant 50 ans et de moins en moins au delà. On constate une bosse à 22 et une bosse plus légère à 68 ans. Pour des âges élevés, l'effet est quasi-inexistant puisque bx est presque nul. C'est explicable grâce au fait que l'amélioration des conditions de vie et de la médecine ont diminué largement la mortalité infantile. De plus, chez les hommes, l'année a un plus grand effet que chez les femmes.

### 2.2.4 Etude l'Ecart absolu des coefficients

```
plot(lca.total$ax-lca.female$ax, main="Ecart avec population totale",
xlab="Age x", ylab=expression(paste(Delta, " ax")), type="l")
lines(x=lca.male$age, y=lca.male$ax-lca.total$ax, main="delta", lty=2)
legend("topright",c("Female","Male"), cex=0.8, lty=1:2)
```

# Ecart avec population totale

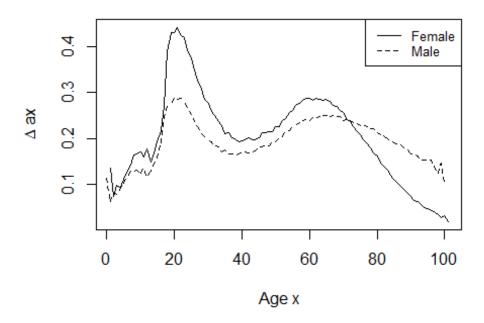


Figure 18: Ecart avec population totale

On constate un plus grand écart de mortalité chez les femmes entre 20 et 30 ans que chez les hommes jusqu'à l'âge de 70 ans. La tendance s'inverse puisqu'au delà, l'écart de mortalité est plus grand pour les hommes.

```
2.2.5 Paramètre k_t
```

```
plot(lca.male$kt, xlab="Year", main="Coef. kt sur données
britanniques",ylab="kt", type="l",ylim=c(-100, 100))
lines(lca.female$year, y=lca.female$kt, main="kt", lty=2)
legend("topright", c("Male", "Female"), cex=0.8, lty=1:2)
```

# Coef. kt sur données britanniques

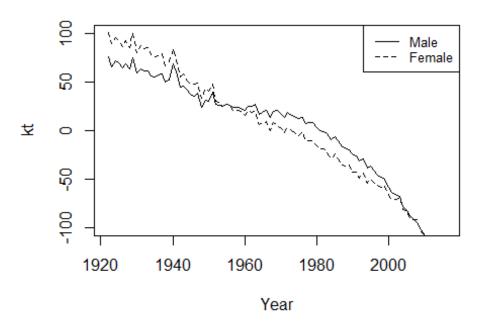


Figure 19: Coef kt sur données britanniques

La valeur kt est en chute constante. On remarque un pic de la valeur de  $k_t$  autour de la deuxième guerre mondiale.

# **2.3** Extrapolation des $(k_t)_t$

Il est possible d'utiliser Kisker ou toute autre méthode de fermeture des tables afin de générer les taux de mortalité à partir des taux lissés aux âges supérieurs à 70. Cependant une recherche poussée à ce sujet n'étant pas parmi nos objectifs académiques, nous nous contenterons de le mentionner.

# Conclusion générale

Malgré la portée limitée du travail académique, il aura permis de générer les notions de base de l'analyse des données de mortalité des pays, afin de pouvoir éventuellement étendre ces connaissances dans le cadre d'un projet réel en assurance vie et gestion des risques.

# **Bibliographie**

http://www.mortality.org/: Mortality tables

Cours d'Actuariat Vie ESPRIT Data Science - Le Mans Université - Chapitre 2

http://medcraveonline.com/BBIJ/BBIJ-06-00158.php: Modeling and Forecasting Norway Mortality Rates using the Lee Carter Model