

Construction des hypothèses Best Estimate biométriques pour un produit  
de Réassurance de dépendance

(Long-Term Care) : Calcul de provision mathématique de rente et  
estimation des IBNR

Hicham BOUKHARSA

# Plan

|      |  |    |
|------|--|----|
| I.   | Description du portefeuille .....      | 05 |
| II.  | Modélisation de la durée.....          | 09 |
|      | 2.1. Estimateur de Kaplan Meier.....   | 11 |
|      | 2.2. Modèle de COX.....                | 18 |
|      | 2.3. Modèle Multi-Etats.....           | 28 |
| III. | Estimation des IBNR.....               | 39 |
|      | 3.1. Méthode de Chain Ladder.....      | 40 |
|      | 3.2. Méthode de Mack Chain Ladder..... | 43 |
|      | 3.3. Méthode de Bootstrap.....         | 46 |
| IV.  | Résultat finale.....                   | 50 |
| V.   | Conclusion.....                        | 54 |

## Modèle pour le calcul de la Provision mathématique de rente

Notre réassureur dérive sa propre table de maintien de dépendance sur la base de l'analyse d'expérience des données brutes. Les taux de maintien se sont détériorés depuis l'année d'occurrence 2012 et il est nécessaire de les mettre à jour dynamiquement pour refléter la tendance la plus récente. La détérioration semble être principalement due au contrôle effectué par le régulateur en 2016.

## Méthode de projection pour le calcul du montant IBNR

IBNR font référence aux sinistres qui se sont produits mais qui n'ont pas encore été signalés au réassureur. Ces sinistres peuvent inclure des événements qui ont eu lieu avant la date d'évaluation, mais qui n'ont pas encore été déclarés ou signalés. Le réassureur doit mettre à jour sa propre vision car les montants des IBNR de la cédante ont historiquement sous-estimé les tendances.

## Remarque et Limitation

- Les taux d'incidence ne sont pas pris en compte car le traité est en "run-off" depuis 2018.
- L'impact de COVID n'est pas anticipé dans cette étude.

# Description du portefeuille

# CARACTÉRISTIQUE DU PRODUIT ÉTUDIÉ ET DU TRAITÉ

## Couverture associée au produit

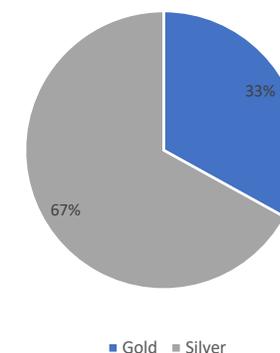
Le portefeuille comprenait deux types de couverture (Silver et Gold) avec des niveaux de prestations différents. Chacun de ces niveaux de couverture ouvrait droit à deux types de prestations forfaitaires : une prestation à domicile et une prestation en institution de santé.

| Couverture | Prestations  |
|------------|--|
| SILVER     | Le dépendant reste à son domicile : \$ 1000/mois   |
|            | Le dépendant se soigne en institution: 40% du cout réel de la prestation plafonné à \$ 1500/mois |
| GOLD       | Le dépendant reste à son domicile : \$ 1200 /mois  |
|            | Le dépendant se soigne en institution : 80% des dépenses dans la limite de \$ 3000 /mois         |

*Silver: Couverture en dépendance total*

*Gold: couverture en dépendance totale et partielle*

Répartition par type de garantie



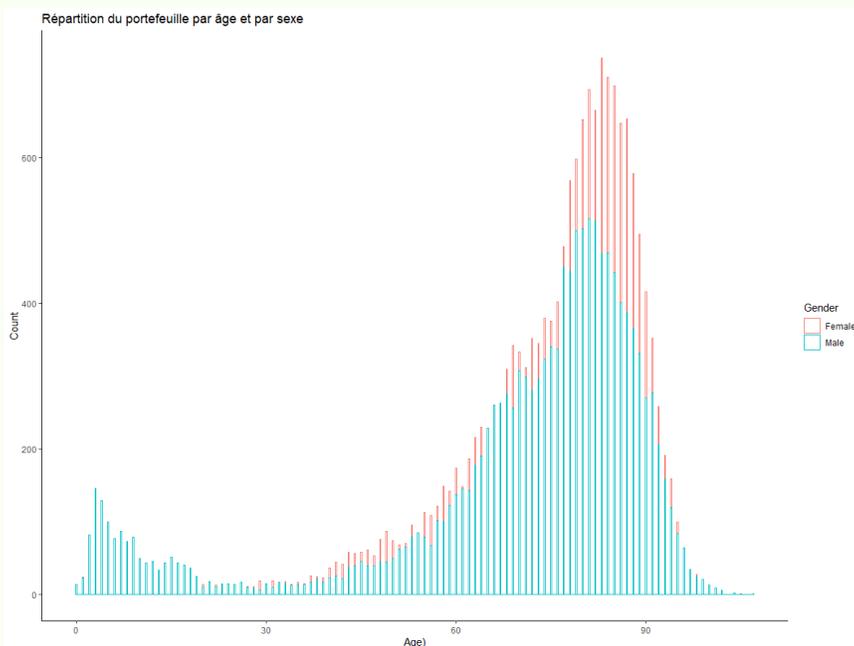
## Les caractéristiques du traité

- Traité en Quote-Part avec un taux de cession de 25%
- Il a été souscrit en 2008 puis en run-off depuis le 31.12.2018.
- Le délai de carence est de 30 jours.
- Le délai maximal de prestation est 60 mois.

# Statistique descriptive du portefeuille (1/2)

## Répartition par âge et par sexe

- Le portefeuille de dépendants total "Silver" est principalement constitué d'assurés entre 61 et 90 ans,
- L'âge moyen est de 73 ans, 70 ans pour les hommes et 75 ans pour les femmes.
- Le plus jeune a 0 an et le plus âgé a 106 ans
- La proportion des femmes est de 53 %, supérieure à celle des hommes (47 %).
- Augmentation nette de dépendants à partir de 75 ans et puis décroissance après l'âge de 85 ans.



Pyramide des âges des assurés du portefeuille étudié par sexe

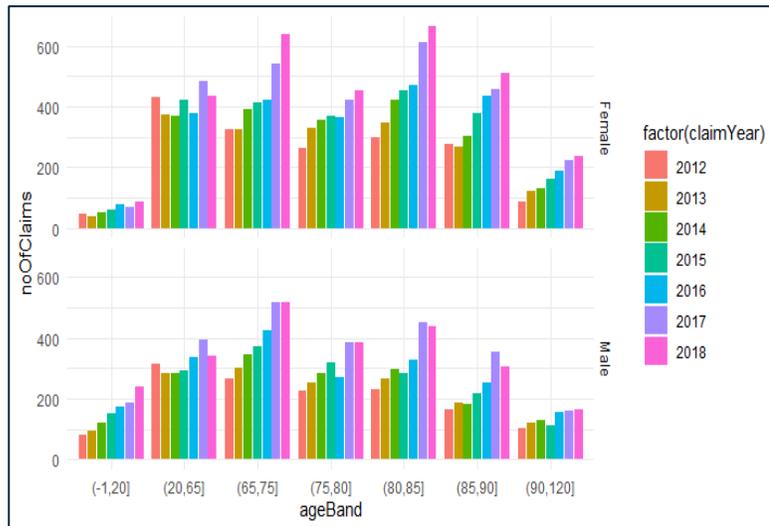
|            | Produit Silver      |                     |        |
|------------|---------------------|---------------------|--------|
|            | Femme en dépendance | Homme en dépendance | Total  |
| Age moyen  | 75,37               | 70,17               | 72,93  |
| Effectif   | 10 857              | 9 593               | 20 450 |
| Proportion | 53,1%               | 46,9%               | 100%   |

Répartition et âge moyen des assurés par sexe pour le produit "Silver"

# Statistique descriptive du portefeuille (2/2)

## Survenance des sinistres

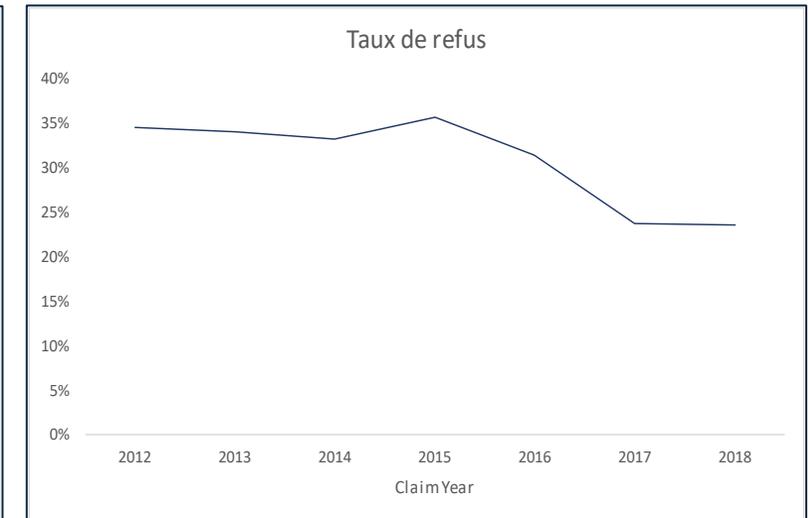
- Le nombre de sinistres déclarés a augmenté régulièrement depuis 2012, atteignant un pic en 2018, cette augmentation concerne les deux sexes.
- Nous remarquons que le nombre de sinistres est en hausse notamment à partir de 2016, cette augmentation s'explique par trois raisons :
  - Hausse de taux d'acceptation (acceptance rate) des sinistres suite à un contrôle effectué par le régulateur local.
  - Augmentation du nombre des sinistres pour les enfants des assurés.
  - Vieillessement du portefeuille des dépendants



Répartition du nombre de sinistres par année de survenance, par sexe et par tranche d'âge



Répartition du nombre de sinistres par année de survenance



Graphique d'évolution du taux de refus par année de survenance.

# Modélisation de la durée

# Analyse de survie

## Description de la survie

- L'analyse de durée porte sur la variable aléatoire  $T$ , représentant le temps écoulé jusqu'à la survenance d'un évènement précis.
- Les modèles de durée présentent certaines particularités :
  - Les données de durée sont engendrées par des variables aléatoires positives.
  - Les données de durée sont souvent incomplètes (observation tronquée et/ou censurée), l'expérience ayant souvent une durée limitée, les assurés ne sont que partiellement observables sur une durée donnée.

## Objectif de l'analyse de survie

- Nous cherchons à estimer :  $S(t) = P(T > t), t \geq 0$ , qui désigne la probabilité que la sortie de l'état de dépendant de l'individu se produise après l'instant  $t$ . Avant l'instant  $t$  l'individu est en état de dépendance.
- Pour la modélisation de la variable de durée  $T$ , différents types de modèles peuvent être utilisés, à savoir :
  - Les modèles paramétriques : Le temps de survie suit une loi de distribution continue et positif (e.g. loi exponentielle, Weibull, Pareto, etc.)
  - Les modèles non-paramétriques : aucune hypothèse a priori n'est faite sur l'allure de la fonction de survie (e.g. estimateur de Kaplan Meier, Nelson-Aalen)
  - Les modèles semi-paramétriques : ils sont basés sur un modèle de base adapté à la cohorte puis différenciés pour chaque sous-groupe (e.g. modèle de Cox)



**Le choix du modèle est orienté par la qualité et la quantité des données à disposition.**



# Estimation non paramétrique: Kaplan Meier

## Cadre théorique

### Estimateur de Kaplan Meier

- L'estimateur de Kaplan-Meier découle de l'idée suivante : être encore en état de dépendance après un instant  $t$ , c'est être en état de dépendance juste avant cet instant  $t$ .

$$\widehat{S}(t) = \prod_{\substack{i=1, \dots, n \\ T_i \leq t}} \left( 1 - \frac{d_{x,i}}{n_{x,i}} \right)$$

- Avec  $T_{(0)} = 0$ . Considérons les notations suivantes :
  - $d_{x,i}$  : le nombre de sorties ayant lieu à l'instant  $T_{x,i}$  pour l'âge à la survenance  $x$
  - $n_{x,i}$  : l'effectif sous risque à l'âge à la survenance  $x$  avant  $T_{x,i}$
- Le principal avantage du modèle de Kaplan-Meier est sa capacité à estimer la probabilité de survie même lorsque certaines observations sont censurées

### Variance de l'estimateur de Kaplan Meier

- L'erreur standard de l'estimateur de Kaplan-Meier est approchée selon la formule de Greenwood:

$$\text{Var}(\widehat{S}(t)) = \widehat{S}(t)^2 \sum_{\substack{i=1, \dots, n \\ T_i \leq t}} \frac{\frac{d_i}{n_i}}{n_i \left( 1 - \frac{d_i}{n_i} \right)} = \widehat{S}(t)^2 \sum_{\substack{i=1, \dots, n \\ T_i \leq t}} \frac{d_i}{n_i (n_i - d_i)}$$

# Estimation non paramétrique: Kaplan Meier

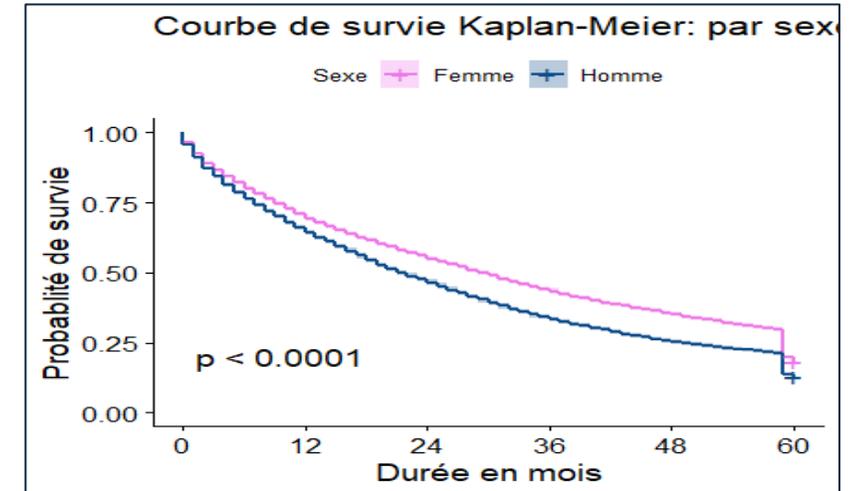
## Segmentation par sexe:

### Courbe de survie brute

- Les courbes de survie pour les hommes et les femmes évoluent de manière quasi-proportionnelle, avec quelques intersections.
- Graphiquement, le taux de maintien est presque le même pour les hommes et les femmes durant les premiers 12 mois.
- Au-delà de 12 mois, les courbes commencent à diverger, indiquant un taux de maintien légèrement inférieur chez les hommes par rapport aux femmes.

### Le test de Log-Rank

- Validation statistique de l'hypothèse de différenciation des lois de maintien entre hommes et femmes se fait via le test de Log Rank
- Dans les sous-populations, une p-value inférieure au seuil de 5 % permet de rejeter  $H_0$ , indiquant une différence significative entre les probabilités des segments Hommes-Femmes



Taux de survie : KM selon le sexe

```
> surv_diff <- survdiff(Surv(claimDurationAdjusted, status) ~ gender, data = data_etude)
> surv_diff
call:
survdiff(formula = Surv(claimDurationAdjusted, status) ~ gender,
  data = data_etude)

      N Observed Expected (O-E)^2/E (O-E)^2/V
gender=Female 10946  8946  9893   90.7    223
gender=Male   9773  8534  7587  118.3    223

Chisq= 223 on 1 degrees of freedom, p= <2e-16
>
```

Sortie R du Test de Log-Rank

# Estimation non paramétrique: Kaplan Meier

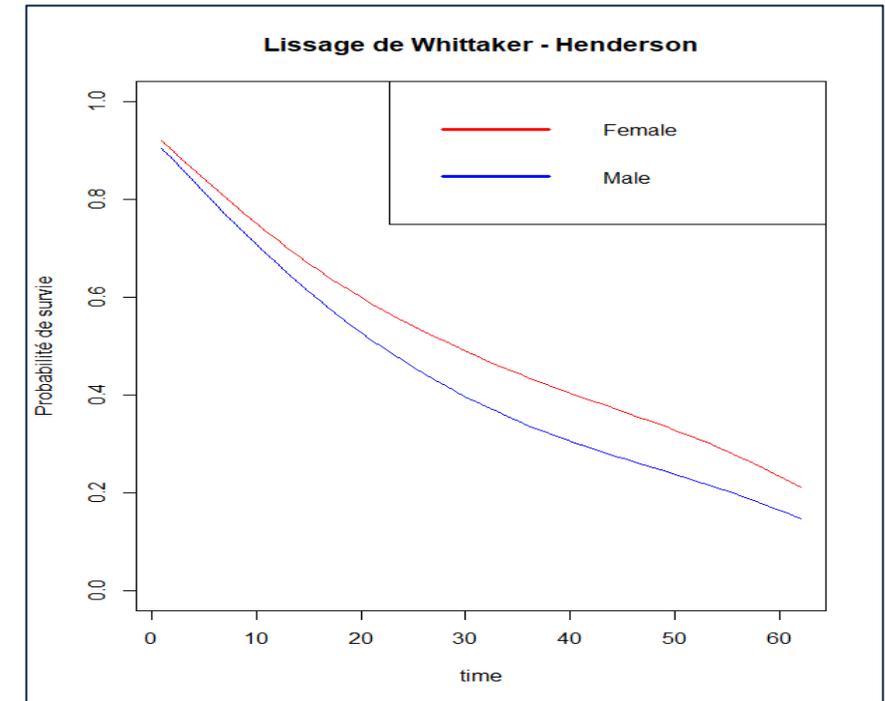
## Segmentation par sexe:

### Courbe de survie lissée

- La courbe des taux bruts présentée précédemment est peu régulière là où il y a peu d'observation. D'où la nécessité de procéder au lissage de ces taux pour pouvoir construire la table de maintien.
- En général, le choix du lissage doit donc dépendre d'un arbitrage entre 2 critères :
  - **La fidélité des taux lissés aux taux bruts**
  - La régularité des taux.

### Méthodes de lissage

- Deux méthodes ont été mises en place, le lissage par la régression de LOESS et le lissage de Whittaker-Henderson (WH).
- Sur la base de la méthode de distance absolue calculée entre les taux bruts et le taux lissés, le lissage de WH permet d'obtenir des taux lissés les plus proche des taux bruts



La courbe de survie : Lissage par Whittaker-Henderson

# Estimation non paramétrique: Kaplan Meier

## Segmentation par tranche d'âge:

### Courbe de survie brute

- L'âge offre des indications utiles sur le maintien en dépendance.
- Junior (0-20 ans) : Dépendance initiale suivie d'une décroissance rapide après 24 mois.
- Adultes (20-65 ans) : Sortie rapide de la dépendance, grâce à une meilleure santé générale.
- Personnes âgées (75-90 ans) : Survie plus courte en dépendance due à une résistance accrue à cet état.
- Très âgés (90 ans et plus) : Forte décroissance de la survie après 30 mois, liée à une faible espérance de vie.

### Le test de Log-Rank

- Une p-value inférieure à 5 % montre une différence significative entre les probabilités de survie des différentes tranches d'âge.

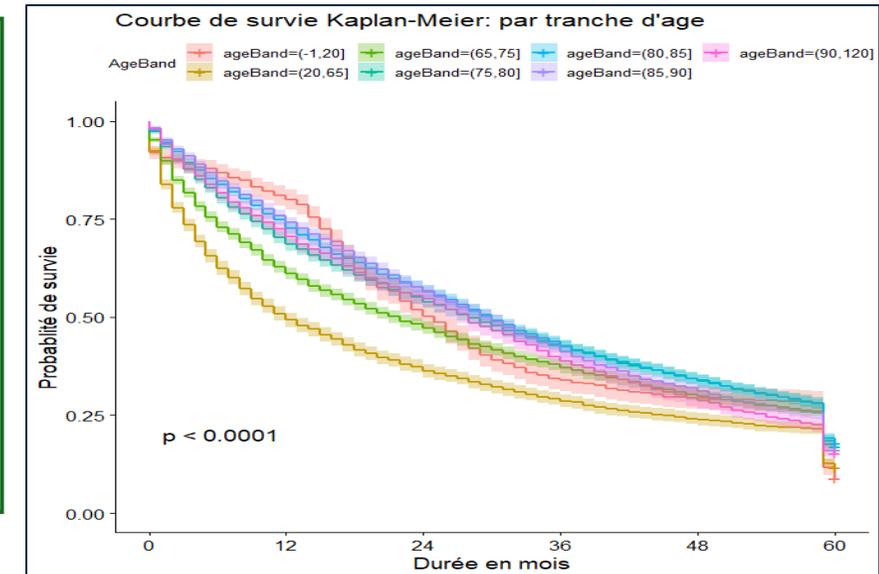


Figure : Taux de survie : KM selon les classes d'âge

```
> surv_diff <- survdiff(Surv(claimDurationAdjusted, status) ~ ageBand, data = data_etude)
> surv_diff
Call:
survdiff(formula = Surv(claimDurationAdjusted, status) ~ ageBand,
  data = data_etude)

      N Observed Expected (O-E)^2/E (O-E)^2/V
ageBand=(-1,20]  960      875      832   2.2391   2.5084
ageBand=(20,65] 3029     2675     2037  199.7939  241.5175
ageBand=(65,75] 3808     3191     3072   4.6101   5.9650
ageBand=(75,80] 3349     2776     2975  13.3530  17.1462
ageBand=(80,85] 4260     3497     3881  37.9821  51.9756
ageBand=(85,90] 3491     2922     3131  13.9186  18.0180
ageBand=(90,120] 1822     1544     1552   0.0424   0.0493

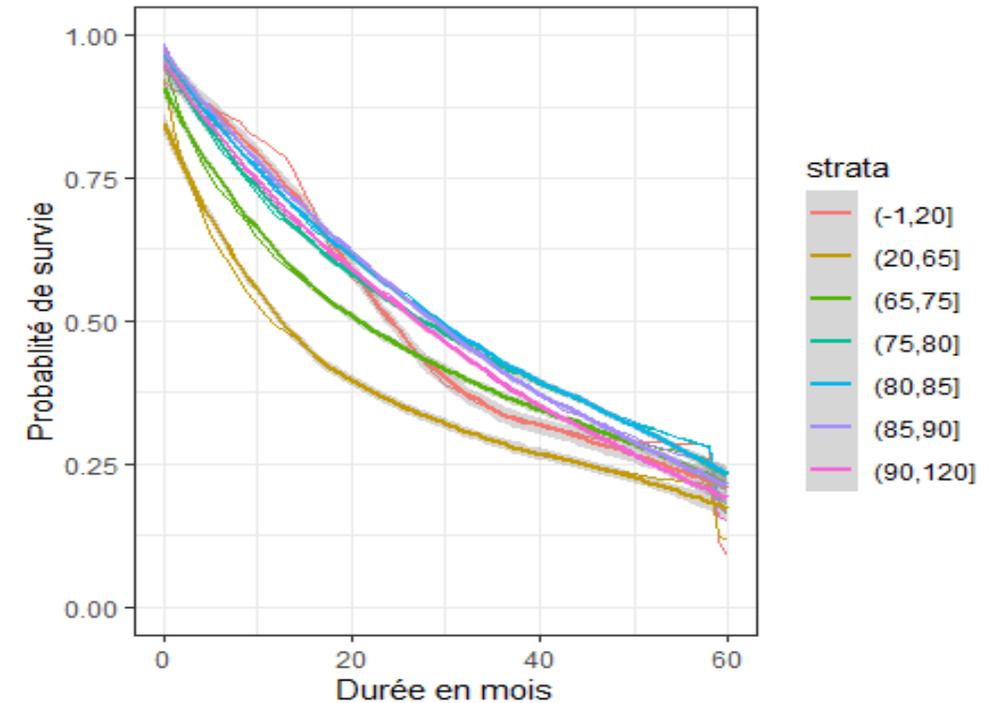
Chisq= 290 on 6 degrees of freedom, p= <math>< 2e-16</math>
```

# Estimation non paramétrique: Kaplan Meier

## Segmentation par tranche d'âge:

### Courbe de survie lissée

- Les taux bruts irréguliers sont lissés pour améliorer leur fiabilité.
- Le lissage par Whittaker-Henderson permet d'obtenir des taux lissés les plus proche des taux bruts selon le critère de la distance absolue.



# Estimation non paramétrique: Kaplan Meier

## Segmentation par sexe et par tranche d'âge:

### Courbe de survie lissée

- Deux lois de maintien en dépendance ont été établies, une pour la population masculine et une pour la féminine, basées sur l'ancienneté en dépendance et la tranche d'âge d'entrée.
- Le lissage de Whittaker – Henderson a été retenu pour lisser les courbes de survie.
- Les courbes de survie diminuent de manière quasi-similaire pour les deux sexes, mais des différences apparaissent selon les tranches d'âge.

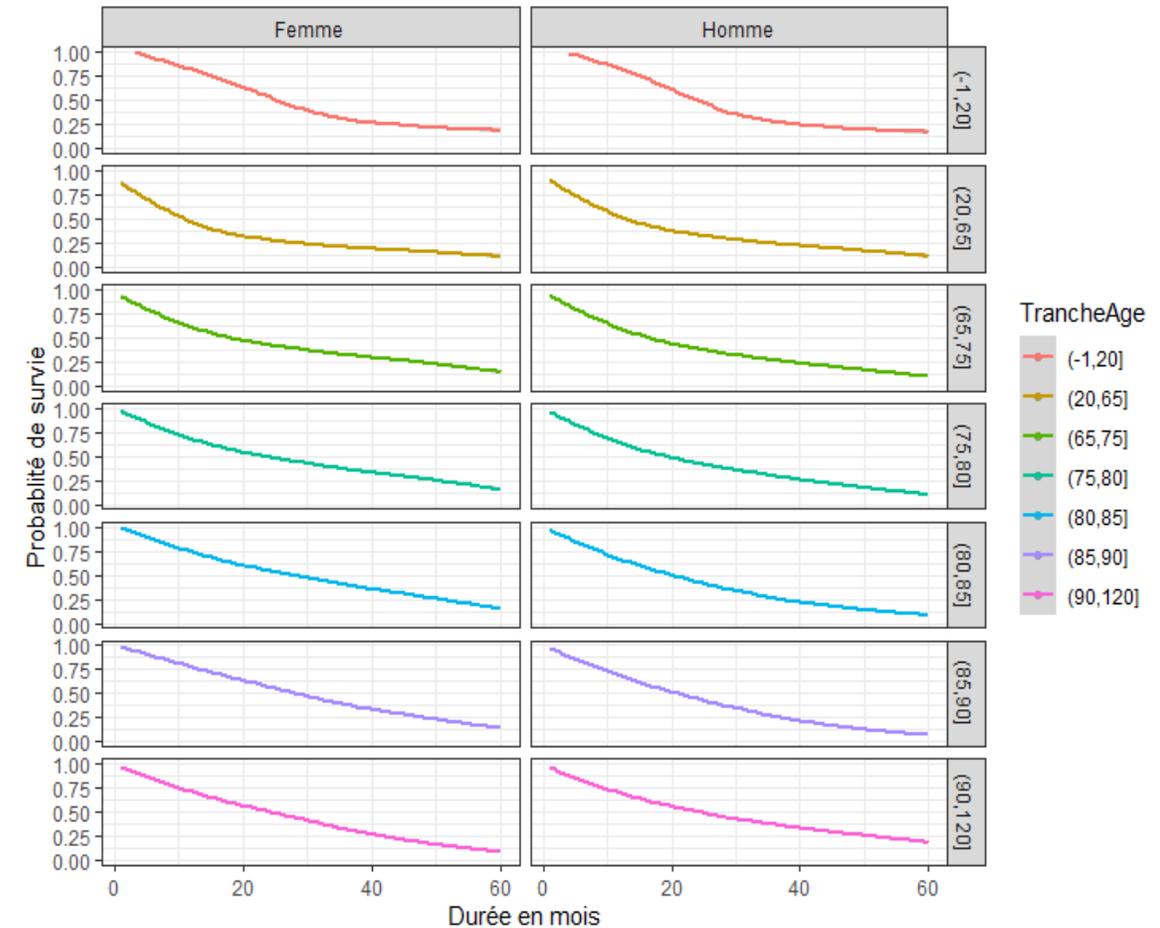


Figure : La courbe de survie lissée par sexe et par tranche d'âge : Lissage par Whittaker-Henderson



# Estimation semi paramétrique: Cox

## Cadre théorique

### Modèle de COX

Le modèle de Cox permet d'exprimer le risque instantané de survenue de l'événement en fonction de l'instant  $t$  et des variables explicatives  $X_j$ .

Le risque instantané (appelé également fonction de hasard) est de la forme :

$$\mu(t, X) = \mu_0(t) \times \exp(\beta' X)$$

Avec :

- $X = (X_1, \dots, X_p)$  un vecteur de variable aléatoires explicatives ;
- $\mu_0$  la fonction de hasard de base commune à tous les individus et qui ne dépend que du temps ;
- $\beta = (\beta_1, \dots, \beta_p) \in \mathbb{R}^p$  paramètre ne dépendant pas du temps : il représente l'effet des covariables sur le risque instantané.

### Hypothèses du modèle de COX

- Les risques sont proportionnels : le rapport des risques instantané d'évènement (hazard rate) de deux individus est proportionnel.
- Il existe une relation log-linéaire entre la fonction de risque instantané et les covariables :

$$\ln \left[ \frac{\mu(t, X)}{\mu_0(t)} \right] = \beta' X$$

# Application semi paramétrique: Cox

## Segmentation par sexe:

### Estimation du modèle de COX

- La variable 'genderMale' est significativement non nulle, indiquant qu'elle est discriminante pour la loi de maintien.
- Impact du coefficient  $\beta$  :
  - Le coefficient  $\beta$  étant positif,  $\exp(\beta)$  est supérieur à 1. Cela indique un risque de dépendance plus élevé chez les femmes par rapport aux hommes.
  - Les hommes ont, en moyenne, 23% moins de chance de rester en dépendance que les femmes.

```

> summary(cox1)
Call:
coxph(formula = surv(duration_adj, status) ~ gender, data = data_etude)

n= 20450, number of events= 17480

              coef exp(coef) se(coef)      z Pr(>|z|)
genderMale 0.21121   1.23517  0.01516 13.93 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

              exp(coef) exp(-coef) lower .95 upper .95
genderMale      1.235      0.8096      1.199      1.272

Concordance= 0.529 (se = 0.002 )
Likelihood ratio test= 193.5  on 1 df,  p=<2e-16
Wald test              = 194.2  on 1 df,  p=<2e-16
Score (logrank) test = 194.9  on 1 df,  p=<2e-16

> print(paste0("Le critère AIC = ",AIC(cox1)))
[1] "Le critère AIC = 315209.096323234"
  
```

Figure : Sortie R de la régression de Cox pour la variable sexe



Ces résultats confirment les premières analyses réalisées avec l'estimateur de Kaplan-Meier.

# Application semi paramétrique: Cox

## Segmentation par sexe:

### Vérification de l'hypothèse de proportionnalité

Pour la variable "Gender", le parallélisme entre les courbes est plus marqué pour les durées supérieures à 22 mois, indiquant une quasi-proportionnalité.

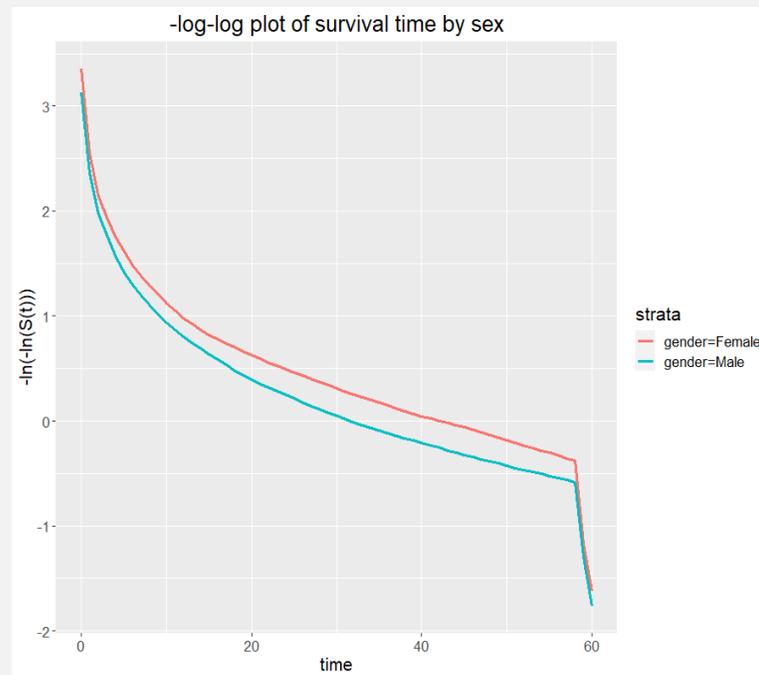


Figure : la courbe LML de la variable sexe

Les résidus Schoenfeld ne montrent pas de tendances systématiques dans le temps et la ligne de régression des résidus est proche de l'horizontale.

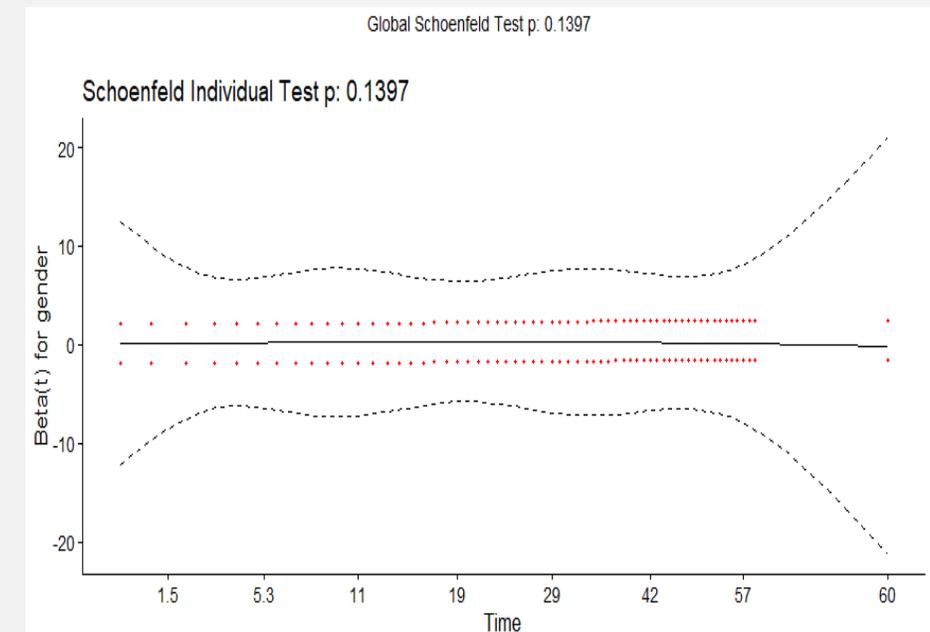


Figure : Test de proportionnalité de la variable sexe à l'aide du test des résidus Schoenfeld

# Application semi paramétrique: Cox

## Segmentation par sexe:

### Courbe de survie du modèle de COX

- Les paramètres permettent de visualiser la fonction de survie pour les deux modalités de la variable "gender".
- **Position relative des courbes** : La courbe de survie des femmes est systématiquement au-dessus de celle des hommes, indiquant que les femmes se maintiennent en dépendance plus longtemps que les hommes.

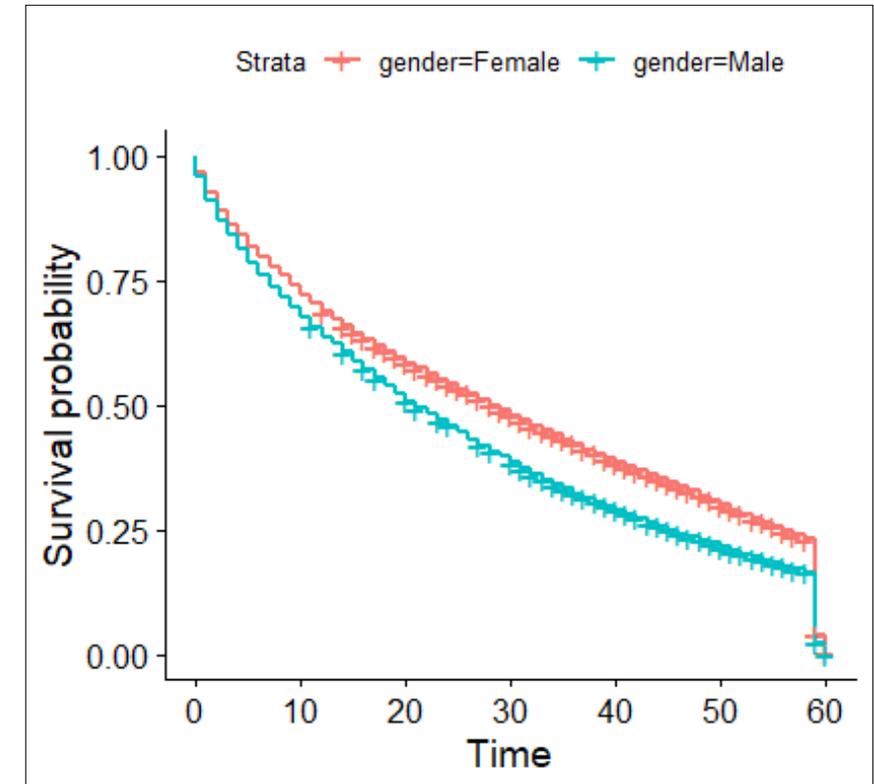


Figure : La courbe de survie pour la variable sexe

# Application semi paramétrique: Cox

Modèle avec 3 variables ("ageBand", "Gender" et "type de service").

## Estimation du modèle de Cox

- La variable "Gender": La dépendance est plus forte chez les femmes que chez les hommes, toutes choses égales par ailleurs
- La variable tranche d'âge : La modalité "ageBand" (90,120] possède l'effet le plus élevé sur le taux de hasard
- La variable "type de service": La modalité "hospitalisation" a un coefficient proche de zéro, ce qui montre que cette variable apporte une explication légère à la fonction de hasard

## Test de proportionnalité

- Il apparaît que la p-value est inférieure à 5% globalement et pour chaque variable prise individuellement.
- Toutes les p-values étant significatives nous acceptons l'hypothèse de non-indépendance entre le temps et les résidus de Schoenfeld.

```
> cox <- coxph(Surv(duration_adj, status) ~ gender + ageBand + Last.Type.of.Services, data = data_etude)
> summary(cox)
Call:
coxph(formula = Surv(duration_adj, status) ~ gender + ageBand +
      Last.Type.of.Services, data = data_etude)

n = 15116, number of events = 12220
(5334 observations effacées parce que manquantes)

              coef exp(coef) se(coef)      z Pr(>|z|)
genderMale    0.21813  1.24375  0.01836 11.881 < 2e-16 ***
ageBand(20,65] 0.82279  2.27684  0.06266 13.131 < 2e-16 ***
ageBand(65,75] 0.62408  1.86653  0.06214 10.043 < 2e-16 ***
ageBand(75,80] 0.53604  1.70923  0.06278  8.538 < 2e-16 ***
ageBand(80,85] 0.51996  1.68195  0.06205  8.380 < 2e-16 ***
ageBand(85,90] 0.56309  1.75609  0.06272  8.978 < 2e-16 ***
ageBand(90,120] 0.69551  2.00472  0.06533 10.646 < 2e-16 ***
Last.Type.of.ServicesHospitalisation 0.07824  1.08138  0.02510  3.117  0.00183 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

              exp(coef) exp(-coef) lower .95 upper .95
genderMale    1.244    0.8040    1.200    1.289
ageBand(20,65] 2.277    0.4392    2.014    2.574
ageBand(65,75] 1.867    0.5358    1.653    2.108
ageBand(75,80] 1.709    0.5851    1.511    1.933
ageBand(80,85] 1.682    0.5945    1.489    1.899
ageBand(85,90] 1.756    0.5694    1.553    1.986
ageBand(90,120] 2.005    0.4988    1.764    2.279
Last.Type.of.ServicesHospitalisation 1.081    0.9247    1.029    1.136

Concordance = 0.574 (se = 0.003 )
Likelihood ratio test = 400 on 8 df, p = < 2e-16
Wald test = 390.9 on 8 df, p = < 2e-16
Score (logrank) test = 395.6 on 8 df, p = < 2e-16

> print(paste0("Le critère AIC = ", AIC(cox)))
[1] "Le critère AIC = 212635.10105571"
```

Figure : Estimation du modèle de Cox pour les variables : "ageBand", "gender" et "type of service"

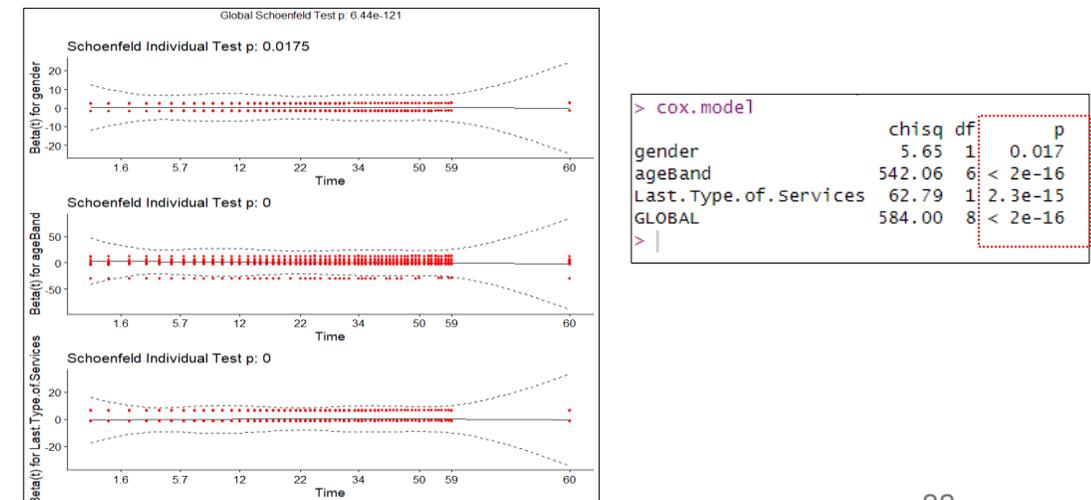


Figure : Représentation graphique des résidus de Schoenfeld

# Application semi paramétrique: Cox

## Stratification par Tranche d'âge de la variable "gender":

### Correction de l'hypothèse de non-proportionnalité:

- Pour remédier au problème de non-proportionnalité, nous pouvons considérer un modèle de **Cox stratifié**. Le principe est d'ajouter des strates au modèle de Cox afin que l'hypothèse de proportionnalité n'ait plus besoin d'être respectée pour la variable "Tranche d'âge".

### Tests de significativité des coefficients

- Le coefficient est significativement différent de 0 avec une p-value < 5%
- Le coefficient  $\beta$  positif montre que les hommes sortent plus rapidement de la dépendance que les femmes.

```
> summary(cox_strata)
Call:
coxph(formula = Surv(duration_adj, status) ~ gender + strata(ageBand),
      data = data_etude)

n= 20450, number of events= 17480

              coef exp(coef) se(coef)      z Pr(>|z|)
genderMale 0.21417  1.23884  0.01539 13.92 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

              exp(coef) exp(-coef) lower .95 upper .95
genderMale    1.239      0.8072    1.202    1.277

Concordance= 0.53 (se = 0.002 )
Likelihood ratio test= 193.1 on 1 df,  p=<2e-16
Wald test              = 193.8 on 1 df,  p=<2e-16
Score (logrank) test = 194.4 on 1 df,  p=<2e-16
```

Figure : Sortie R pour le modèle de Cox stratifié

# Application semi paramétrique: Cox

Stratification par Tranche d'âge de la variable "gender":

## Représentation graphique de la courbe de survie:

- Les fonctions de survie correspondant à chaque groupe peuvent être représentées graphiquement.

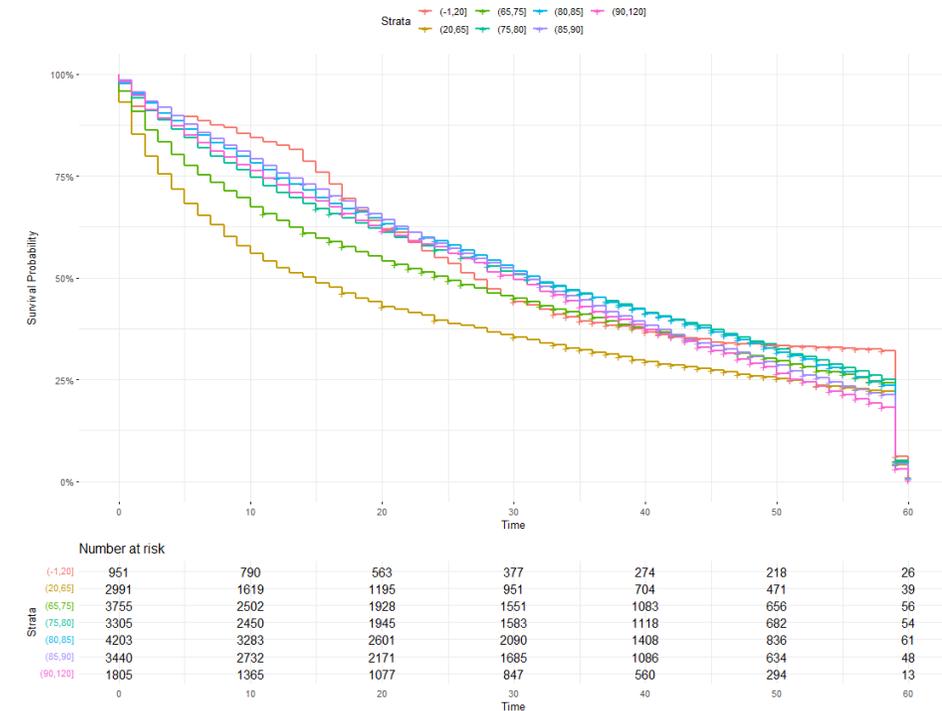


Figure : La courbe de survie du modèle Cox-stratifié

# Application semi paramétrique: Cox

## Stratification par Tranche d'âge de la variable "gender":

### Modèle de Cox stratifié avec interaction

- Les modèles stratifiés ne différencient pas l'effet des covariables entre les groupes, rendant nécessaire l'inclusion d'une interaction entre le genre et les groupes d'âge.
- Les variables utilisées dans le modèle sont significativement discriminantes pour la loi de maintien, avec un seuil de risque de 5 %
- Le risque de dépendance est inférieur chez les hommes par rapport aux femmes, avec des réductions variant de 18 % à 44,6 % selon les tranches d'âge supérieures à 65 ans. Cependant, pour les 20-65 ans, les hommes ont tendance à rester dépendants plus longtemps que les femmes.

```
> cox_strata <- coxph(Surv(duration_adj, status) ~ gender+gender*strata(ageBand), data = data_etude)
> summary(cox_strata)
Call:
coxph(formula = Surv(duration_adj, status) ~ gender + gender *
      strata(ageBand), data = data_etude)

n = 20450, number of events = 17480

              coef exp(coef) se(coef)      z Pr(>|z|)
genderMale      0.06515  1.06732  0.07526  11.881 < 2e-16 ***
genderMale:strata(ageBand)(20,65] -0.11103  0.89491  0.08463  13.131 < 2e-16 ***
genderMale:strata(ageBand)(65,75]  0.10078  1.10603  0.08322  8.978 < 2e-16 ***
genderMale:strata(ageBand)(75,80]  0.19611  1.21666  0.08436  2.325  0.020083 **
genderMale:strata(ageBand)(80,85]  0.23588  1.26602  0.08264  2.854  0.004314 ***
genderMale:strata(ageBand)(85,90]  0.30364  1.35478  0.08415  3.608  0.000308 ***
genderMale:strata(ageBand)(90,120]  0.20841  1.23172  0.09091  2.292  0.021881 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

              exp(coef) exp(-coef) lower .95 upper .95
genderMale      1.0673  0.9369  0.9209  1.237
genderMale:strata(ageBand)(20,65]  0.8949  1.1174  0.7581  1.056
genderMale:strata(ageBand)(65,75]  1.1060  0.9041  0.9396  1.302
genderMale:strata(ageBand)(75,80]  1.2167  0.8219  1.0313  1.435
genderMale:strata(ageBand)(80,85]  1.2660  0.7899  1.0767  1.489
genderMale:strata(ageBand)(85,90]  1.3548  0.7381  1.1488  1.598
genderMale:strata(ageBand)(90,120]  1.2317  0.8119  1.0307  1.472

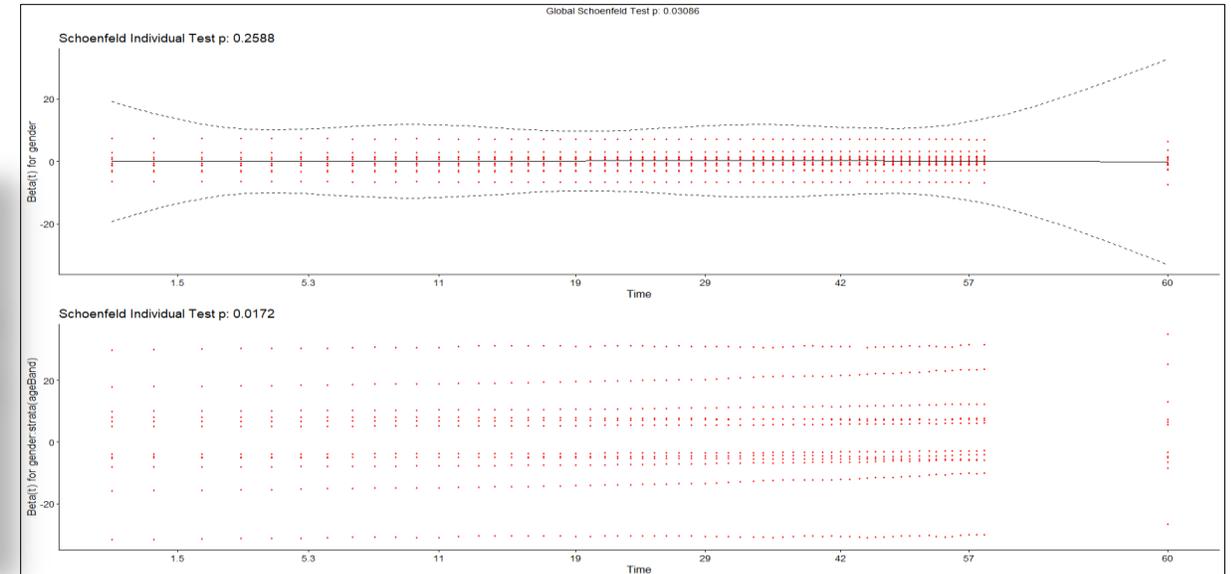
Concordance = 0.533 (se = 0.002 )
Likelihood ratio test = 270 on 7 df,  p=<2e-16
Wald test              = 273.5 on 7 df,  p=<2e-16
Score (logrank) test = 275.6 on 7 df,  p=<2e-16
```

Figure : Estimation du modèle de Cox stratifié avec interaction selon la variable "ageband"

## Stratification par Tranche d'âge de la variable "gender":

### Vérification de l'hypothèse de proportionnalité: Résidus de Schoenfeld

**Diagnostic :** L'hypothèse des risques proportionnels est désormais vérifiée, et de plus nous jugeons une amélioration conséquente du modèle selon le critère AIC.



```
> print(paste0("Le critère AIC = ",AIC(cox_strata)))
[1] "Le critère AIC = 249568.795891712"
> cox.model <- cox.zph(cox_strata)
> cox.model
```

|                        | chisq | df | p      |
|------------------------|-------|----|--------|
| gender                 | 1.28  | 1  | 0.259  |
| gender:strata(ageBand) | 15.42 | 6  | 0.017  |
| GLOBAL                 | 15.43 | 7  | 0.1385 |

Figure Sortie R du test de vérification de l'hypothèse de proportionnalité et représentation graphique des résidus Schoenfeld



# Modèle multi-états

## Cadre théorique: Modèle de Markov

- Propriété de Markov : l'information sur les états précédents est résumée par l'état présent  
 Les intensités de transition  $\mu$  peuvent dépendre de deux échelles de temps :
  - La durée du suivi  $t$
  - Le temps de séjour
- On distingue plusieurs types de modèles multi-états :
  - Markovien homogène (MH) : les intensités de transition sont constantes ;
  - Markovien non-homogène (MNH) : les intensités de transition dépendent du temps
  - Semi-markovien homogène (SMH) : les intensités de transition dépendent de la durée passée dans l'état
  - Semi-markovien non-homogène (SMNH) : les intensités de transition dépendent de la durée passée dans l'état et du temps

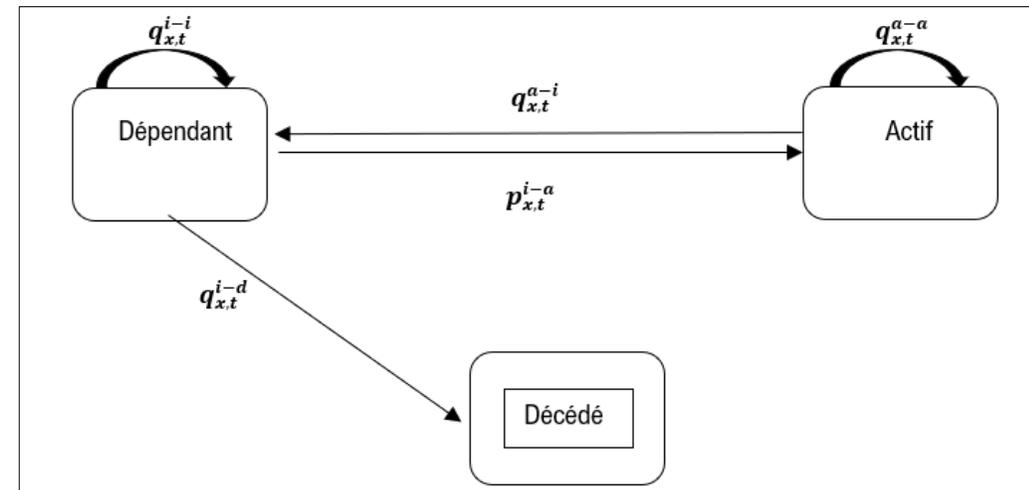


Figure: Etats relatifs au risque dépendance du produit "Silver"

# Modèle multi-états

## Cadre théorique: Modèle de Markov

- $(X_t)_{t \geq 0}$  un processus de Markov a temps continu et à espace d'états fini  $S = \{e_1, e_2, \dots, e_m\}$ . Le processus  $(X_t, \mathcal{F}_t)_{t \geq 0}$  est markovien s'il satisfait à la propriété de Markov, c'est-à-dire s'il vérifie l'égalité suivante pour  $0 \leq s \leq t$  et pour tout  $e \in S$   $P(X_t = e / \mathcal{F}_t) = P(X_t = e / X_s)$
- Intensités de transition entre les états  $Q(t)$ , les éléments de cette matrice sont définis par :
- $q_{hj}(t) = \delta_{h,j} \mu_{hh}(t) + (1 - \delta_{h,j}) \mu_{hj}(t)$  avec  $\mu_{hj}(t) = \lim_{\Delta t \rightarrow 0} \frac{p_{hj}(t, t+\Delta t)}{\Delta t}$  pour  $h \neq j$  et  $\mu_{hh}(t) = -\sum_{j \neq h} \mu_{hj}(t)$

## Cas de modèle Markov homogène

Pour tout  $0 \leq s \leq t$ ,  $P(s, t) = P(0, t - s)$  on notera alors simplement  $P(t - s) = P(0, t - s)$ , on dit que les probabilités de transition sont stationnaires, ainsi :  $P(x) = \exp(Q \times x)$

## Cas de modèle Markov non homogène

Les intensités de transition dépendent de la durée du suivi : les probabilités de transitions pour tout  $0 \leq s \leq t$ :  $P(x) = \exp\left(\int_s^t Q(u) du\right)$

# Modèle multi-états

## Modèle Markovien homogène.

### Application du modèle Markov

- L'estimation de la matrice d'intensité permet de calculer les probabilités de transition et par conséquent tracer une courbe de survie pour chaque état.
- Le temps moyen de séjour dans un état valide ou dépendant est estimé par  $1/q_{rr}$ , où  $q_{rr}$  est le terme diagonal de la matrice d'intensité de transitions estimée.
- Les courbes de survie décroissent exponentiellement, plus vite chez les assurés en état "valide".

```
> sojourn.msm(cav.msm)
      estimates      SE      L      U
State 1  48.296 0.5797 47.173 49.446
State 2   1.731 0.0206  1.691  1.772
> |
```

Figure : Sortie R de la matrice d'intensité de transition estimée et temps moyen de séjour dans un état

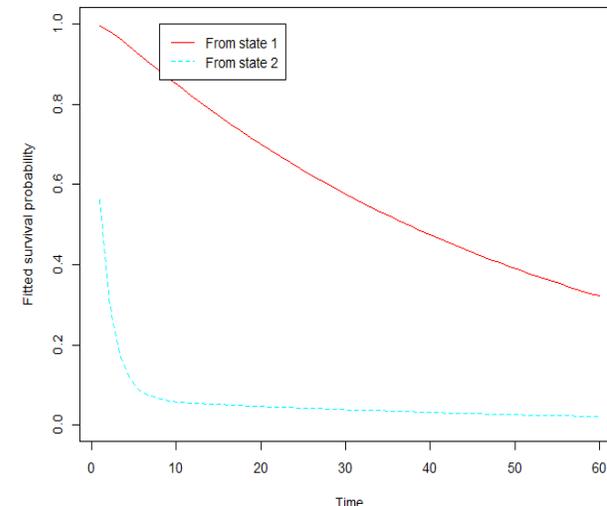


Figure : La fonction de survie pour les deux états "dépendant" et "valide"

## Modèle multi-états

### Modèle Markovien homogène.

#### Validation du modèle Markov

- Pour valider le modèle, les valeurs prédites peuvent être comparées aux valeurs observées.
- La validation montre que le modèle correspond bien aux données pour les états "dépendant" et "décès". Cependant, pour l'état "sain", il sous-estime la prévalence avant 10 mois et la surestime ensuite, probablement à cause du faible nombre d'observations.

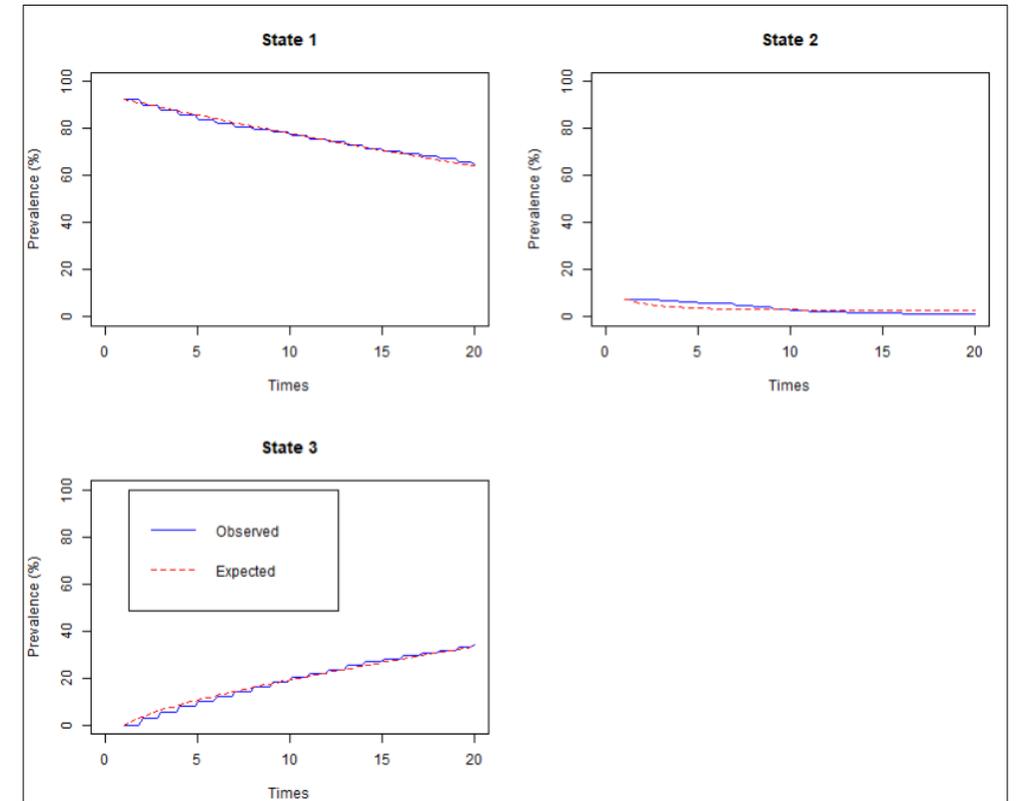


Figure : Sortie R de la fonction de "prévalence" pour les différents statuts

En général, Il apparaît que l'hypothèse d'un modèle markovien homogène est valable

## Modèle multi-états

### Modèle Markovien homogène.

Partant du portefeuille initial constitué des personnes en état de dépendance, soit l'état  $h$  et on génère l'état suivant,  $j$  à l'aide des probabilités estimées des passages  $(p_{h,j})_{h \neq j}$



A partir de l'état initial  $h$  et l'état suivant  $j$ , nous générons le temps de passage correspondant par une réalisation d'une loi exponentielle simple.



Le dernier état visité ne s'agit pas d'un état de sortie (décès), nous répétons les étapes 1 puis 2.

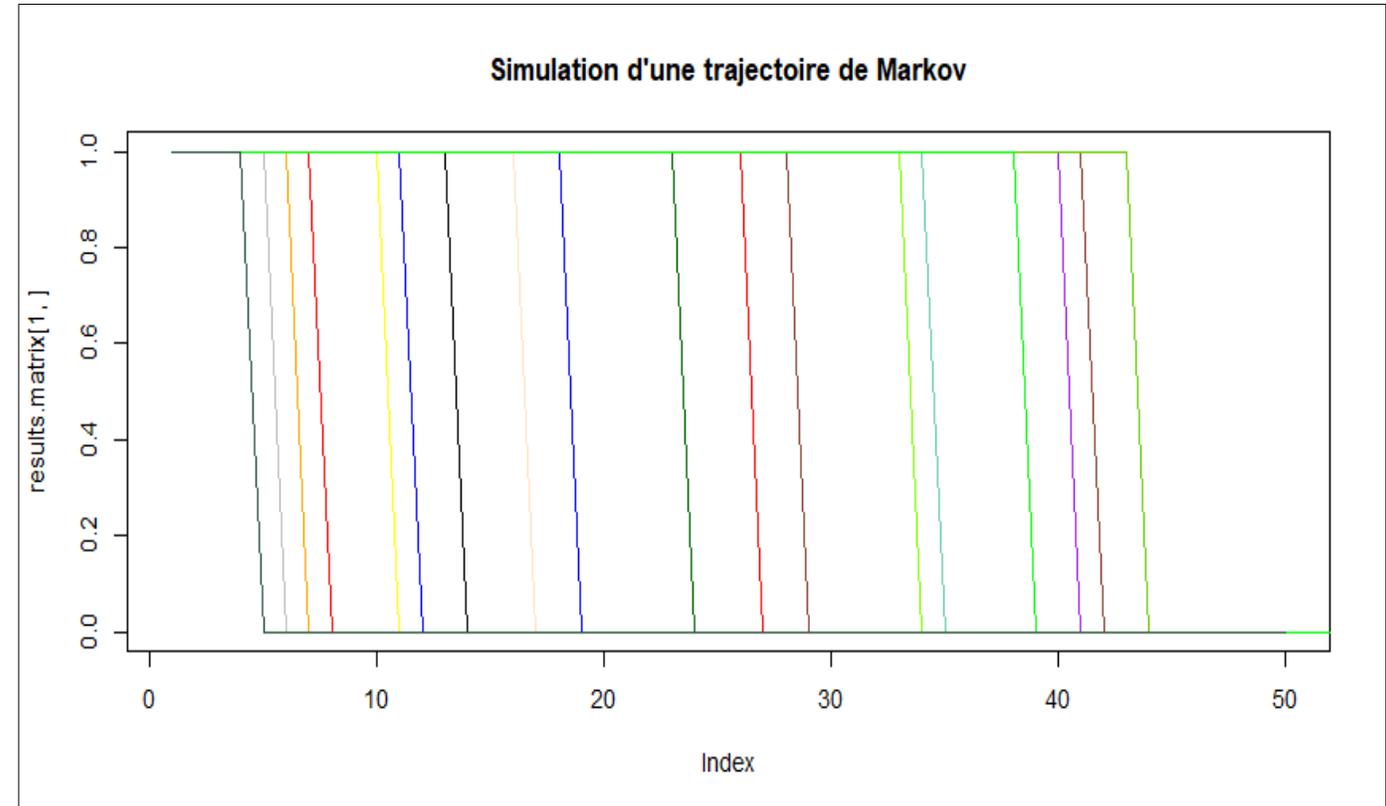


Figure : Simulation de 25 trajectoires de l'état "dépendant"

# Modèle multi-états

## Modèle Markovien localement homogène.

### Principe du modèle Markov localement homogène

Une modélisation des chaînes de Markov localement homogène, présume que une homogénéité dans le temps n'est vérifiable que par des morceaux (modèle avec des intensités constantes par morceaux).

### Processus markovien homogène par sexe:

Les courbes de survies relatives à la variable sexe est conforme aux résultats précédents : la dépendance chez les femmes (durée de maintien) est supérieure à celles des hommes.

```
> qmatrix.msm(cavgender.msm, covariates=list(sexe="1")) #Female
      State 1          State 2          State 3
State 1 -0.01807 (-0.01868,-0.01748)  0.01807 ( 0.01748, 0.01868)  0
State 2  0.04127 ( 0.03601, 0.04729) -0.57040 (-0.58928,-0.55213)  0.52914 ( 0.51170, 0.54717)
State 3  0                               0                               0

> qmatrix.msm(cavgender.msm, covariates=list(sexe="0")) #Male
      State 1          State 2          State 3
State 1 -0.01807 (-0.01868,-0.01748)  0.01807 ( 0.01748, 0.01868)  0
State 2  0.04127 ( 0.03601, 0.04729) -0.57040 (-0.58928,-0.55213)  0.52914 ( 0.51170, 0.54717)
State 3  0                               0                               0

> sojourn.msm(cavgender.msm, covariates=list(sexe="1")) #Female
      estimates      SE      L      U
State 1  55.343 0.93552 53.540 57.208
State 2   1.753 0.02912  1.697  1.811
> sojourn.msm(cavgender.msm, covariates=list(sexe="0")) #Male
      estimates      SE      L      U
State 1  40.57 0.69251 39.237 41.952
State 2   1.41 0.02392  1.364  1.458
> |
```

Figure : Le temps moyen de séjour dans un état en fonction du sexe

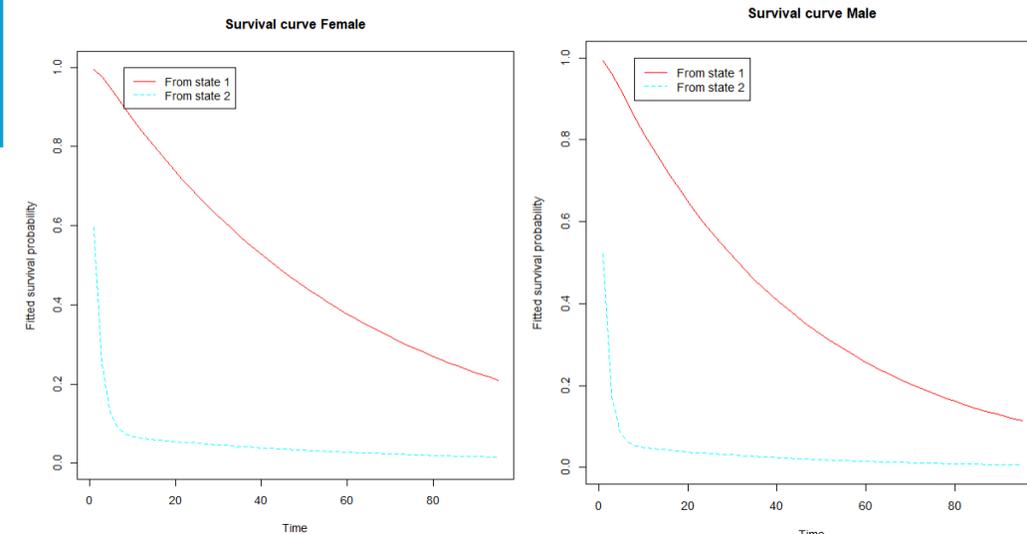


Figure : Fonction de survie par sexe en fonction de l'état

# Modèle multi-états

## Modèle Markovien localement homogène.

### Validation du modèle Markov localement homogène

Le modèle prédit globalement bien la prévalence par sexe, il sur-estime légèrement la prévalence pour l'état dépendant chez les femmes et l'état décès chez les hommes, et il sous-estime la prévalence pour l'état dépendant chez les hommes et l'état décès chez les femmes.

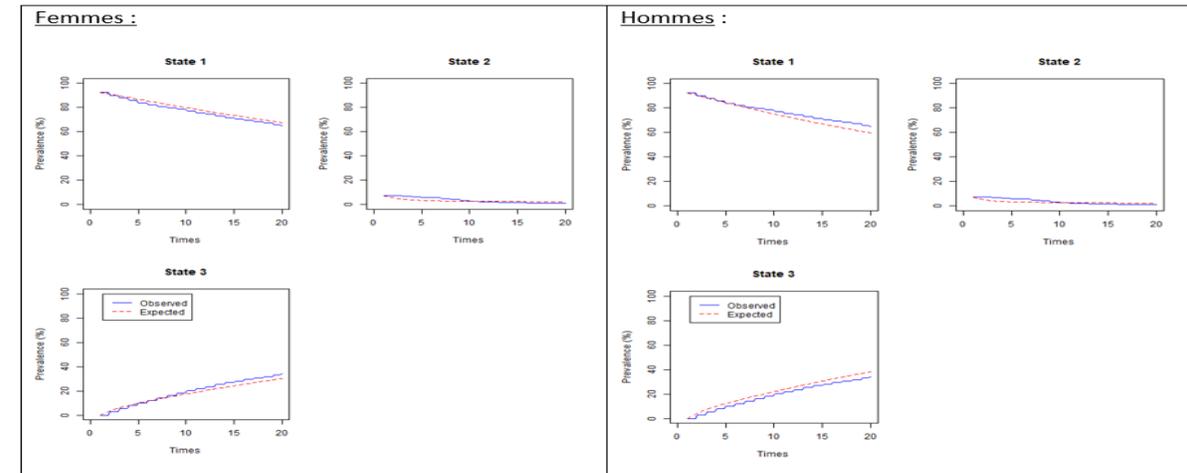


Figure : Sortie R de la fonction de "prévalence" en fonction du sexe pour les différents statuts

# Modèle multi-états

## Modèle Semi-Markovien.

### Cadre théorique: Modèle semi-Markov homogène

On définit le modèle semi-markovien homogène comme suit :

- $S_k$  correspond, pour tout  $k \in \mathbb{N}^*$ , à la date du  $k^{\text{ème}}$  saut du processus  $(X_t)_{t \geq 0}$   $S_k = \inf\{t > S_{k-1} / X_t \neq X_{S_{k-1}}\}$
- $J_k$  correspond, pour tout  $k \in \mathbb{N}^*$ , à l'état du processus  $(X_t)_{t \geq 0}$  entre les dates  $S_k$  et  $S_{k+1}$
- Le processus  $(S_k; J_k)_{k \geq 0}$  permet de notifier les temps de saut ainsi que l'état occupé entre chaque saut, et fournit une représentation complète de la trajectoire suivie par l'état de l'individu.
- Le processus  $(X_t)_{t \geq 0}$  est semi-markovien s'il vérifie l'égalité suivante, pour  $k \in \mathbb{N}^*$  et pour tout  $C \in [0, +\infty[ \times S$ 

$$P((S_{k+1}; J_{k+1}) \in C / (S_l; J_l); l = 1, \dots, k) = P((S_{k+1}; J_{k+1}) \in C / (S_k; J_k))$$
- Dans le processus semi-Markovien homogène les intensités de transition dépendent de la durée passée dans l'état et de l'état occupé.

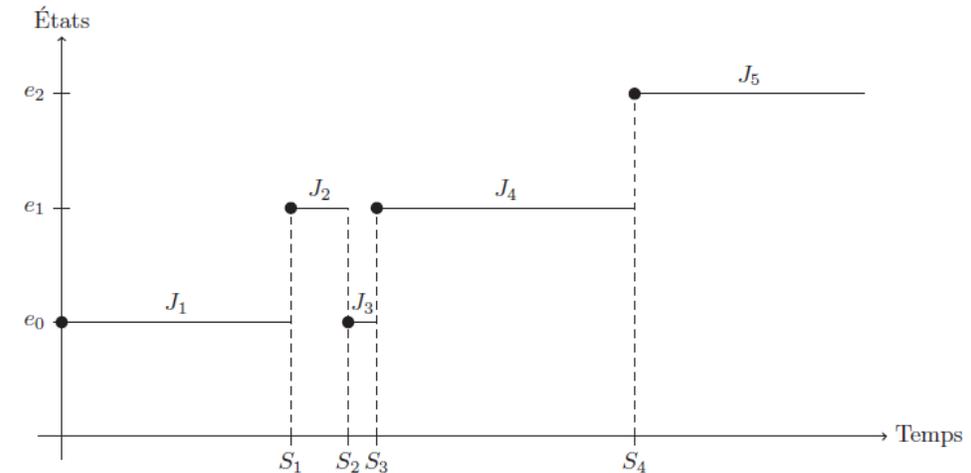


Figure : Exemple de chemin parcouru au cours de la durée de vie d'un individu pour un modèle multi-états quelconque à 3 états  $\{e_0, e_1, e_2\}$ .

# Modèle multi-états

## Modèle Semi-Markovien.

### Modèle semi Markov localement homogène par sexe

- Estimation du temps de séjour à l'aide de la loi de Weibull sera réalisé par le biais du modèle à temps de vie accélérée AFT (Accelerated Failure Time)
- Les modèles à temps de vie accélérée s'écrivent :  $\log(T) = \beta_1 X_1 + \dots + \beta_p X_p + \varepsilon$  où la fonction de survie prend la forme :  $S(t/X) = S_0(t * \exp(-\beta'X))$
- Les fonctions de survie sont discernables. Les femmes ont une probabilité de survie plus élevée que celle des hommes en cohérence avec les modèles précédents.

```
> summary(weibull1)
Call:
survreg(formula = Surv(duration_adj1, status) ~ gender, data = data_weibull,
        dist = "w")
              Value Std. Error      z      p
(Intercept)  3.58031    0.00967  370.2 <2e-16
genderMale   -0.18229    0.01382  -13.2 <2e-16
Log(scale)   -0.09149    0.00644  -14.2 <2e-16

Scale= 0.913

Weibull distribution
Loglik(model)= -78186   Loglik(intercept only)= -78272
      Chisq= 173.4 on 1 degrees of freedom, p= 1.3e-39
Number of Newton-Raphson Iterations: 5
n= 20450
```

Figure : Sortie R de la régression AFT

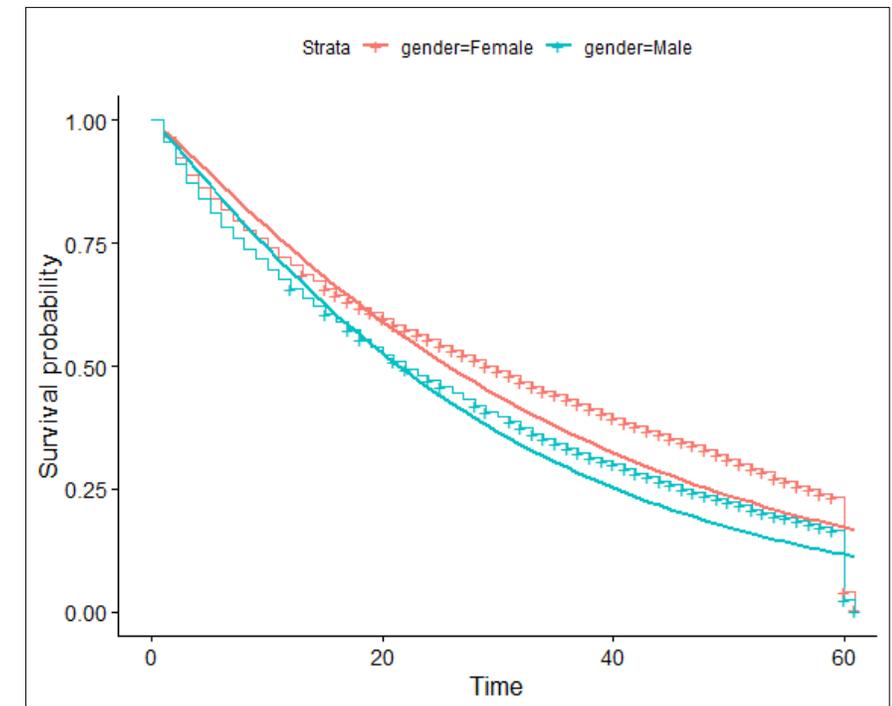


Figure : la fonction de survie estimé par la régression d'AFT en fonction du sexe

# Modèle multi-états

## Modèle Semi-Markovien.

L'évolution de l'état de l'assuré d'un mois à un autre s'effectue de la manière suivante ; soit  $u$  une réalisation de variable aléatoire uniforme :

A partir de l'état initial de dépendance  $i_1$  nous générons l'état suivant  $i_2$ , si  $u < p^{dep,dep}$ , l'assuré reste en dépendance

À partir de l'état initial  $i_1$  et l'état suivant  $i_2$ , le temps de transition est généré en utilisant une loi de Weibull si  $i_2$ . Les paramètres de cette loi sont estimés à l'aide du modèle AFT.

Si le dernier état visité n'est pas la mort, nous répétons les étapes 2 puis 3 jusqu'au décès de l'individu.

```

> beweibull1<- flexsurvreg(Surv(duration_adj1,status)~gender, data=data_weibull1, dist="weibull")
> beweibull1
Call:
flexsurvreg(formula = Surv(duration_adj1, status) ~ gender, data = data_weibull1,
            dist = "weibull")

Estimates:
      data mean  est      L95%      U95%      se      exp(est)  L95%      U95%
shape      NA  1.09581  1.08206  1.10972  0.00706      NA      NA      NA
scale      NA 35.88461 35.21078 36.57134  0.34707      NA      NA      NA
genderMale 0.47330 -0.18229 -0.20938 -0.15520  0.01382  0.83336  0.81109  0.85625

N = 20450, Events: 17480, Censored: 2970
Total time at risk: 569087
Log-likelihood = -78186, df = 3
AIC = 156377
  
```

Figure : Sortie R de l'estimation des paramètres de la loi de Weibull.

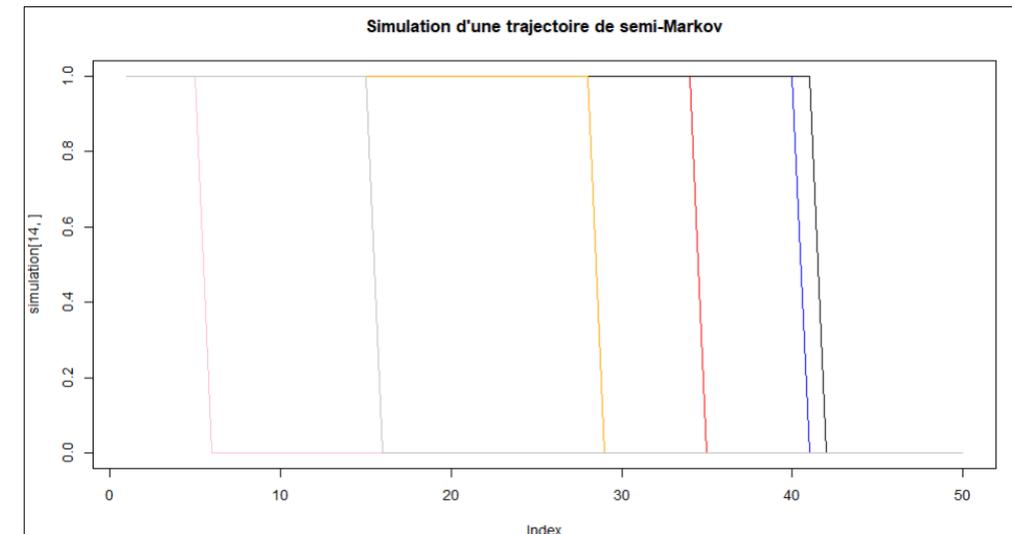


Figure : Simulation semi-markovienne de 5 trajectoires de l'état "dépendant"

# Calcul des IBNR



# Modèle Chain Ladder

## Cadre théorique: Méthode de Chain Ladder.

### Cadre théorique: Méthode de Chain Ladder

- La méthode Chain Ladder est une méthode déterministe fréquemment utilisée car facile à mettre en œuvre  
L'idée de cette méthode est de supposer que la liquidation future est similaire à la liquidation passée. Les hypothèses du modèle sont les suivantes :
- le nombre de sinistres incrémentaux sont indépendants par mois de survenance,
- les mois de développement sont les variables explicatives du comportement du nombre de sinistres futurs.

- On définit le coefficient de passage par :  $\hat{f}_j = \frac{\sum_{i=0}^{n-j+1} C_{i,j+1}}{\sum_{i=0}^{n-j+1} C_{i,j}}$

Grâce à ces facteurs, nous pouvons estimer :

- Le nombre de sinistres ultimes par exercice de survenance  $\widehat{C}_{i,n} = C_{i,n-i} \times \prod_{j=n-i}^{n-1} \hat{f}_j$
- Le nombre de sinistres tardifs par exercice de survenance  $\widehat{IBNR}_i = \widehat{C}_{i,n} - C_{i,n-i}$

# Modèle Chain Ladder

## Application: Méthode de Chain Ladder.

### Vérification de l'hypothèse de linéarité du Chain Ladder

- L'alignement des couples  $(C_{i,j}, C_{i,j+1})$ : Afin de vérifier cette hypothèse nous allons examiner l'existence d'une relation linéaire entre le nombre de sinistres cumulés d'un mois de déroulement à l'autre.

### Résultat de la méthode de Chain Ladder

- Les résultats des estimations issues de la méthode de Chain Ladder standard sont résumés dans le tableau suivant.
- Le nombre de sinistre tardif que le réassureur estime recevoir s'élève à 125 sinistres

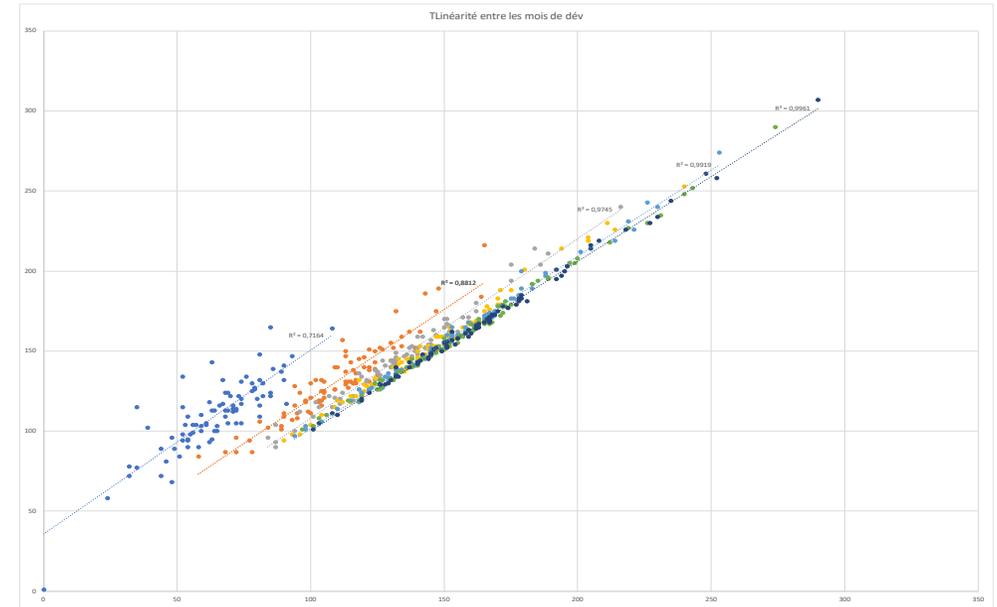


Figure : C-C plot de nombre de sinistres cumulés

|                                   | 2016  | 2017  | 2018  | Total        |
|-----------------------------------|-------|-------|-------|--------------|
| <b>Nombre de sinistre observé</b> | 2 382 | 2 860 | 3 637 | <b>8 879</b> |
| <b>Nbre de sinistre Ultime</b>    | 2 383 | 2 866 | 3 755 | <b>9 004</b> |
| <b>IBNR nombre</b>                | 1     | 7     | 117   | <b>125</b>   |



# Modèle Mack Chain Ladder

## Cadre théorique: Méthode de Chain Ladder.

### Hypothèse 1

- Les mois de survenance des sinistres sont indépendants les uns des autres, c'est-à-dire,  $C_{i,j}$  et  $C_{k,j}$  sont indépendants si  $i \neq k$ .

### Hypothèse 2

- Suppose que le passage d'un mois de développement à l'autre est décrit en termes d'espérance

$$\forall i = 1, \dots, n, \forall j = 1, \dots, n - 1, E(C_{i,j+1}/C_{i,1}, C_{i,2}, \dots, C_{i,j}) = C_{i,j} \times f_j$$

### Hypothèse 3

- Suppose que la variance des sinistres est proportionnelle à son espérance

$$\forall i = 1, \dots, n, \forall j = 1, \dots, n - 1, \text{il existe un paramètre } \sigma_j$$

$$\text{Var}(C_{i,j+1}/C_{i,1}, C_{i,2}, \dots, C_{i,j}) = C_{i,j} \times \sigma_j^2$$

# Modèle Mack Chain Ladder

## Hypothèses du Modèle de Mack Chain Ladder

### Vérification de l'hypothèse H1:

Afin de vérifier l'indépendance des mois de survenance, nous étudions l'influence de l'effet calendaire sur les facteurs de développement individuels.

### Vérification de l'hypothèse H2

Afin de vérifier l'hypothèse d'espérance, les couples  $(C_{i,j}, C_{i,j+1})_{i=0, \dots, n-j-1}$  doivent être sensiblement alignés par une droite passant par l'origine.

### Vérification de l'Hypothèse 3

la vérification de l'hypothèse de variance se fera en analysant le caractère aléatoire des résidus à l'aide de diagrammes de dispersion

### Résultat IBNR

- A l'aide du vecteur des estimations de l'écart type, et sous l'hypothèse de normalité des résidus du nombre estimé, l'intervalle de confiance à 95% du nombre d'IBNR est  $[95.24 ; 154.83]$  avec un nombre moyen IBNR= 125.04

# Méthode de Bootstrap

## Application de la méthode de Bootstrap

$X_{i,j} \sim ODP(m_{i,j}, \phi)$  par hypothèse et les coefficients  $\hat{m}_{i,j}$  et  $\hat{\phi}_j$  sont estimés



Calcul des résidus de Pearson de la loi de Poisson (dans ce cas c'est  $\frac{X_{i,j}}{\phi_j}$  qui suit une loi de Poisson et non  $X_{i,j}$ )

$$r_{i,j} = \frac{X_{i,j} - \hat{m}_{i,j}}{\sqrt{\hat{\phi}_j \cdot \hat{m}_{i,j}}}$$


Ajustement des résidus  $r_{i,j}^{adj} = \sqrt{\frac{n}{\frac{1}{2}n(n+1) - 2n + 1}} \times r_{i,j}$



Pour chaque nouveau triangle de résidus:

- Calcul de la table des pseudo-nombre incrémentés en inversant la formule des résidus
- Par la méthode de Chain Ladder nous calculons à nouveau la partie inférieure de triangle
- Estimation des nombres finaux et déduction des IBNR à constituer



Les résidus sont rééchantillonnés de façon aléatoire et avec remise

# Modèle Bootstrap

## Application de la méthode de Bootstrap

### Application de la méthode de Bootstrap

- Nombre de simulations 5000
- La modélisation ODP par Bootstrap donne une moyenne de 121,1 sinistres IBNR avec un quantile à 95 % de 152 sinistres.
- Les résidus sont supposés suivre une loi normale centrée-réduite, hypothèse est validée par le test de Kolmogorov-Smirnov à un seuil de 5 %.

```
> ks.test(Residuals_boot,"pnorm")
Asymptotic one-sample Kolmogorov-Smirnov test
data: Residuals_boot
D = 0.16967, p-value = 0.3168
alternative hypothesis: two-sided
```

Figure : Sortie R du test de Kolmogorov-Smirnov

```
> Bot
BootChainLadder(Triangle = Tri_Nbre, R = nbsim, process.distr = "od.pois")
Latest Mean Ultimate Mean IBNR IBNR.S.E IBNR 75% IBNR 95%
1 193 193 0.000 0.000 0 0 0
2 156 156 0.000 0.000 0 0 0
3 203 203 0.000 0.000 0 0 0
4 173 173 0.000 0.000 0 0 0
5 177 177 0.000 0.000 0 0 0
6 214 214 0.000 0.000 0 0 0
7 274 274 0.237 0.643 0 0 1
8 226 226 0.187 0.571 0 0 1
9 167 167 0.138 0.481 0 0 1
10 139 139 0.120 0.439 0 0 1
11 241 241 0.323 0.754 1 2 2
12 219 219 0.291 0.698 0 0 2
13 281 281 0.368 0.806 1 2 2
14 204 204 0.268 0.677 0 2 2
15 205 205 0.271 0.676 0 2 2
16 195 195 0.264 0.665 0 2 2
17 246 246 0.318 0.759 1 2 2
18 244 244 0.389 0.828 1 2 2
19 248 248 0.454 0.883 1 2 2
20 253 254 0.505 0.937 1 2 2
21 210 211 0.574 0.955 1 2 2
22 219 220 0.673 1.018 1 3 3
23 273 274 0.990 1.268 2 3 3
24 281 282 1.283 1.381 2 4 4
25 264 266 1.534 1.509 2 4 4
26 226 228 1.780 1.559 3 5 5
27 307 310 2.742 1.939 4 6 6
28 274 278 3.573 2.202 5 8 8
29 297 303 6.369 2.882 8 11 11
30 243 250 6.759 2.955 9 12 12
31 323 334 10.596 3.712 13 17 17
32 285 295 10.483 3.724 13 17 17
33 179 186 7.352 3.056 9 13 13
34 298 311 13.217 4.176 16 20 20
35 331 347 16.407 4.675 19 24 24
36 610 643 32.595 6.778 37 44 44

Totals
Latest: 8,878.0
Mean Ultimate: 8,999.1
Mean IBNR: 121.1
IBNR.S.E: 18.4
Total IBNR 75%: 133.0
Total IBNR 95%: 152.0
```

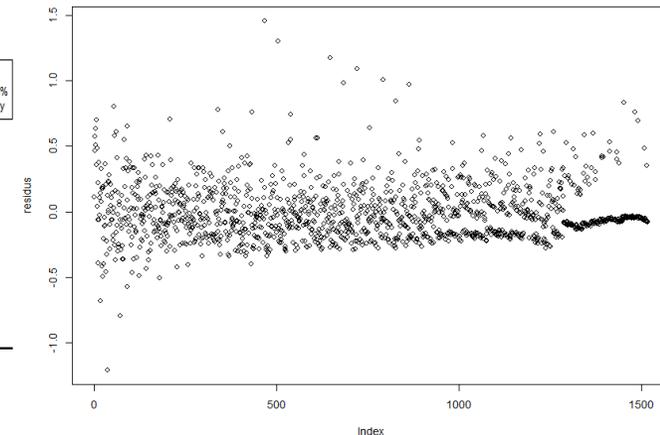
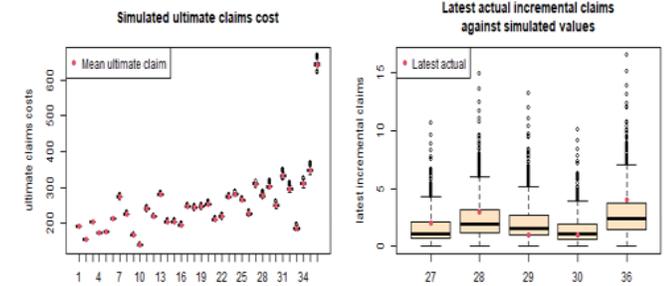
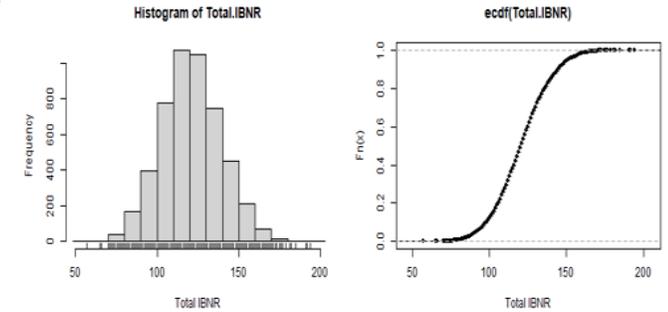
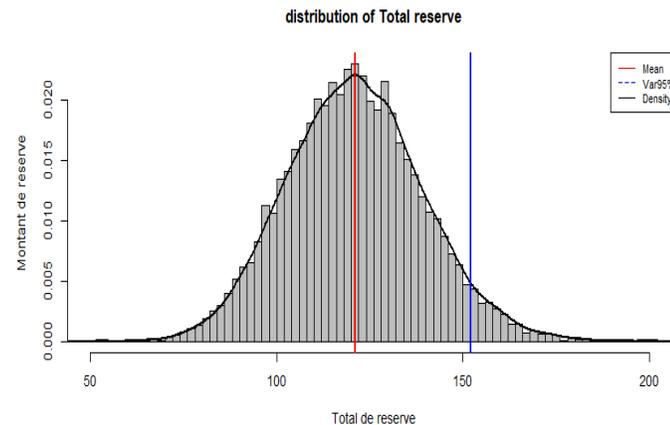


Figure : Sortie R du résultat de l'estimation de nombre IBNR avec la méthode de Bootstrap.

# Comparaison des modèles

## Comparaison entre les méthodes

- Les méthodes de Mack Chain Ladder et Chain ladder donnent des résultats similaires, mais Bootstrap montre un faible nombre de sinistres tardifs ; une approche prudentielle favorise la méthode Mack pour sa prudence.
- La méthode Bootstrap affiche une variabilité plus élevée que Mack, dont les résidus semblent plus aléatoires et moins volatiles autour de la moyenne.

|  | Méthode de Chain Ladder | Mack Chain ladder | Bootstrap |
|--|-------------------------|-------------------|-----------|
| <b>Moyenne de la distribution</b>                      | 125.04                  | 125.04            | 121.1     |
| <b>Ecart type</b>                                      |                         | 15.2              | 18.4      |
| <b>VaR de niveau 99.5%</b>                             |                         | 166               | 173       |
| <b>Coefficient de volatilité</b>                       |                         | 12.2%             | 15.2%     |
| <b>Coefficient de comparaisons <math>VaR/BE</math></b> |                         | 32.76%            | 42.98%    |

Tableau de comparaison entre les différentes méthodes d'estimation des IBNR

# Résultat de calcul IBNR

## La mise en place de la méthode de calcul des IBNR

- L'estimation du nombre IBNR nécessaire s'établit à 125 nombres. La répartition est effectuée par tranche d'âge et sexe selon la structure du portefeuille.
- Nous comptons 114 sinistres en attentes qu'on ajoute au nombre IBNR des sinistres avec une probabilité d'acceptation établie à 85%.

Ainsi le nombre total de sinistre tardif est  $125 + 114 * 85\% = 222$  *sinistres*. La répartition se présente comme suit :

|              | IBNR + sinistre en attente |            |           |             |
|--------------|----------------------------|------------|-----------|-------------|
| Sexe         | ageBand                    | % effectif | Age moyen | Nombre IBNR |
| <b>Femme</b> | (-1,20]                    | 2,95%      | 8,6       | 3           |
|              | (20,65]                    | 15,33%     | 54,2      | 18          |
|              | (65,75]                    | 17,51%     | 71,0      | 21          |
|              | (75,80]                    | 15,62%     | 78,3      | 18          |
|              | (80,85]                    | 21,51%     | 83,1      | 25          |
|              | (85,90]                    | 17,87%     | 87,8      | 21          |
|              | (90,120]                   | 9,22%      | 93,2      | 11          |
|              | <b>Total Femme</b>         |            | 100,00%   | 75,4        |
| <b>Homme</b> | (-1,20]                    | 8,77%      | 8,0       | 9           |
|              | (20,65]                    | 15,83%     | 54,7      | 16          |
|              | (65,75]                    | 20,61%     | 70,7      | 21          |
|              | (75,80]                    | 15,71%     | 78,2      | 16          |
|              | (80,85]                    | 17,39%     | 82,9      | 18          |
|              | (85,90]                    | 13,34%     | 87,9      | 14          |
|              | (90,120]                   | 8,35%      | 93,4      | 9           |
|              | <b>Total Homme</b>         |            | 100%      | 70,2        |

Tableau : Répartition du nombre IBNR et nombre de sinistre en attente par sexe et par tranche d'âge

# Calcul du provisionnement technique en dépendance

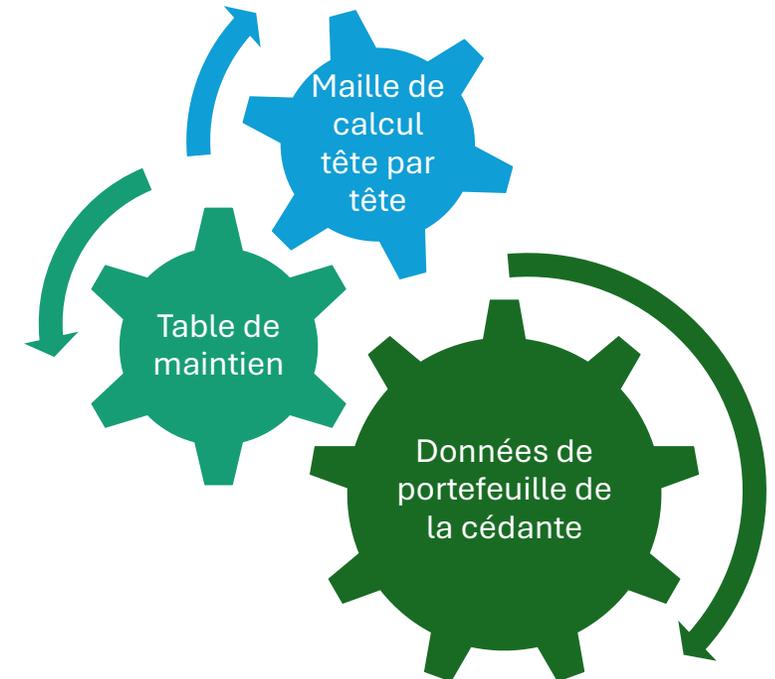
# Le calcul de la PM de rente en dépendance

## Le calcul de la PM de rente en dépendance

- La Provision mathématiques de rente correspond à la provision pour rente en cours de service
- Grâce aux données de la cédante, la provision mathématique est calculée comme suit :  $PM = a^{t,s} \times R$  Où  $a^{t,s} = \sum_{k=1}^T v_k^k \prod_{j=0}^{k-1} {}_{j|1}p_x^{d,s}$

Où:

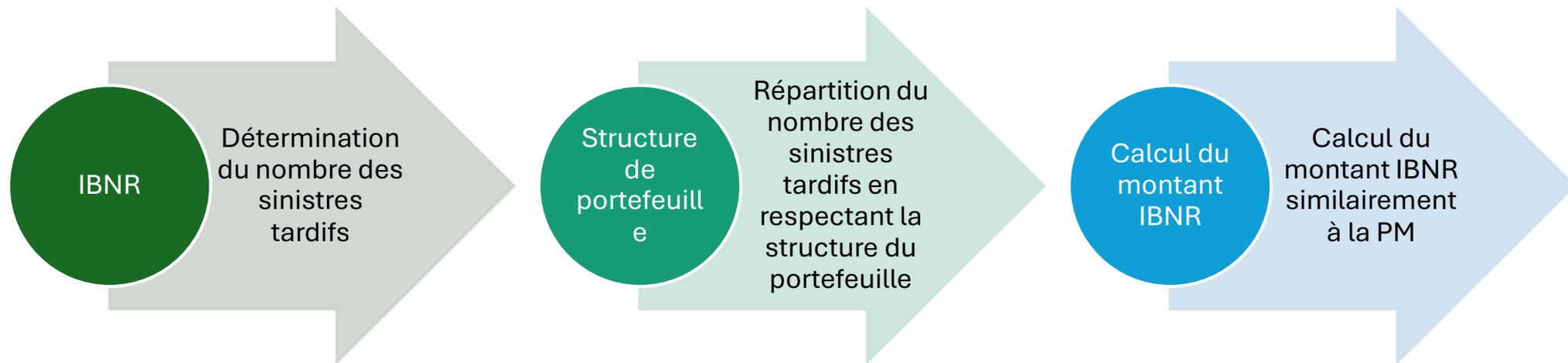
- ${}_{k|1}p_x^{d,s}$ : la probabilité qu'un individu de sexe S, dépendant totalement d'âge  $x$  reste dans le même état pour les âges compris entre  $x + k$  et  $x + k + 1$ .
- $v = 1 / (i + 1)$  où  $i$  est le taux d'intérêt
- $R$  montant de l'annuité de la garantie.
- La réglementation locale du réassureur oblige d'utiliser la courbe des taux "locked-in"



# Le calcul des IBNR.

## Démarche retenue pour le calcul des IBNR

- La méthode adoptée consiste à projeter le nombre de sinistres tardifs (IBNR) en suivant le rythme des sinistres survenus et connus. Pour cela, le nombre de sinistres tardifs est réparti en respectant la structure du portefeuille pour ensuite utiliser les tables de maintien retenus pour le calcul de la PM afin de projeter dans l'avenir les dates possibles de paiement.



# Résultat final du montant des provisions techniques.

## Le résultat de l'estimation des provisions techniques

- Le montant de 45 millions USD représente la somme nécessaire aujourd'hui pour couvrir les engagements de la compagnie de réassurance jusqu'à l'extinction du groupe fermé. Le montant de 2,25 millions USD, correspondant à la Provision pour Risque Adverse (PAD), vise à compenser l'incertitude liée à l'estimation des flux futurs pour les risques non financiers cette provision pour déviation défavorable est calculée de manière à garantir que l'évaluation des engagements soit suffisante à hauteur de 75 %.

| Amount in m USD              | Cedant account | Reinsurance's Evaluation |
|------------------------------|----------------|--------------------------|
| In-payment                   | 30.78          | 37.5                     |
| IBNR + OS pnding             | 2.25           | 2.14                     |
| <b>Total reserve</b>         | <b>33.03</b>   | <b>39.64</b>             |
| Paid claims                  | 5.26           | 5.26                     |
| <b>Total BE</b>              | <b>38.29</b>   | <b>44.9</b>              |
| <b>Total BE + PAD(*)</b>     | <b>40.20</b>   | <b>47.15</b>             |
| <b>Deficit(-)/Surplus(+)</b> |                | <b>-6.94</b>             |

Résultat total de montant de provision technique

# Conclusion

## Conclusion générale

**Objectif** : Développement d'un modèle jugé complet pour le calcul des provisions techniques définies comme la somme de la provision mathématique et des IBNR d'un produit en dépendance,

**Modélisation de la provision mathématique** : Trois approches pour déterminer les lois de maintien.

- - **Kaplan-Meier (lissage Whittaker-Henderson)**,
- - Modèle de Cox stratifié,
- - Modèle multi-états markovien.
- 
- **Estimation IBNR** : Application des méthodes de Triangulation:
  - - Méthode Chain-Ladder,
  - - **Mack Chain Ladder** ,
  - - Méthode de Bootstrap.

**Provisions techniques** : Établies à 47.15 M USD incluant la provision pour déviation adverse.

**Limites** : Hypothèses d'indépendance et absence de tables de référence restreignant la validation.

**Perspectives** : Le risque de dépendance étant récent et évolutif, cette étude n'est pas définitive et pourra être révisée en fonction des évolutions réglementaires et juridiques à venir

Merci