Équipe 01

Travail pratique II

Alexandre Simard ()

Carl Villeneuve Lepage ()

Catherine Laflamme ()

David Baril ()

Modèles linéaires en actuariat

ACT-2003

Travial présenté à Olivier Côté

Faculté des sciences et de génie

Université Laval

Le xx décembre 2021

[Premier projet : Modélisation de la fréquence des consultations médicales pour l’assurance maladie 4](#_Toc771716905)

[Faits saillants 4](#_Toc2071664520)

[1. Loi de Poisson avec lien canonique 4](#_Toc2019316423)

[a) Estimation du maximum de vraisemblance de β et estimation des écart-types 4](#_Toc167035744)

[b) Élaboration du modèle courant 4](#_Toc665982341)

[c) Pearson, déviance et AIC 5](#_Toc267574996)

[d) Estimations ponctuelles sur ID 1075 à 1084 6](#_Toc248608538)

[e) Estimation ponctuelle et intervalle de confiance 95% P(Yi =0) 7](#_Toc1963246934)

[f) Graphiques des fonctions de masse de probabilité 8](#_Toc695296423)

[g) Graphiques des résidus de Pearson, d’Anscombe et de déviance 10](#_Toc1477046103)

[2. Loi binomial négative avec lien logarithmique 12](#_Toc841716890)

[a) EMV de β et estimation des écart-type 12](#_Toc283097583)

[b) Élaboration du modèle courant 12](#_Toc2061819784)

[c) Comparaison des résultats avec le modèle Poisson 12](#_Toc228180104)

[d) Pearson et AIC 12](#_Toc2067797901)

[e) Estimation ponctuelle sur ID 1075 à 1084 12](#_Toc2079833519)

[f) Estimation ponctuelle et intervalle de confiance 95% P(Yi =0) 12](#_Toc10739460)

[g) Test statistique 12](#_Toc1221029980)

[3. Modèle recommandé 13](#_Toc137277042)

[a) Présentation du modèle 13](#_Toc980228291)

[b) Estimation des paramètres et des écart-types 13](#_Toc1886249160)

[c) Statistiques du modèle 13](#_Toc886461503)

[d) Graphiques des fonctions de masse de probabilité 13](#_Toc799340710)

[e) Caractéristique de risque pour VaR0.99 13](#_Toc1332014319)

[Second projet : Taux d’abandon au renouvellement de polices d’assurance automobile 14](#_Toc366614640)

[Faits saillants 14](#_Toc903537198)

[1. Modèle binomial 14](#_Toc2061864644)

[a) Expression de probabilité d’abandon 14](#_Toc1101163758)

[b) Fonction de lien 14](#_Toc1216114416)

[c) Estimation des paramètres et des écart-types 14](#_Toc206770566)

[d) Transformation et traitement des variables explicatives et des interactions 14](#_Toc1023519801)

[e) Sélection des variables par la déviance 14](#_Toc98198446)

[f) Courbe ROC 14](#_Toc674533400)

[g) Maximisation de la sensivité et de la spécifité 14](#_Toc977235159)

[2. Prévisions 15](#_Toc715950405)

[a) Estimation ponctuelle 15](#_Toc615293354)

[b) Interprétation et explication des résultats 15](#_Toc972110263)

[c) Tableau de mauvaise classification 15](#_Toc208916364)

[d) Prévisions et explications 15](#_Toc1294214950)

# Premier projet : Modélisation de la fréquence des consultations médicales pour l’assurance maladie

## Faits saillants

## 1. Loi de Poisson avec lien canonique

### a) Estimation du maximum de vraisemblance de β et estimation des écart-types

Nous avons d’abord élaboré un modèle incluant toutes les variables du jeu de données sans en modifier le contenu. Voici, dans le tableau \_, les résultats des estimations du maximum de vraisemblance et des écart-types pour les variables explicatives du modèle complet. La variable réponse est “docvis”.

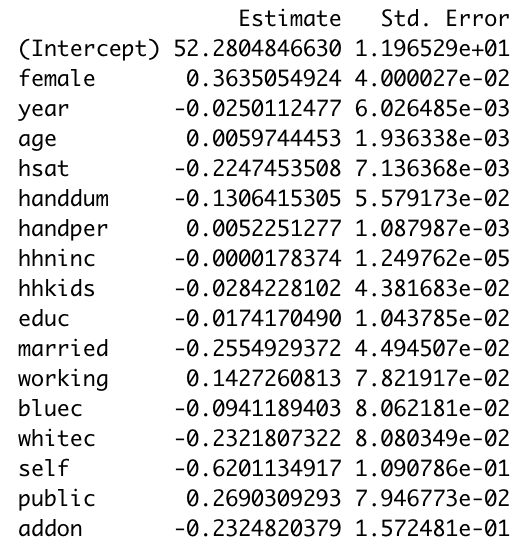


Tableau \_ : EMV et écart-types pour β

### b) Élaboration du modèle courant

La sélection de variables a été faite avec l’analyse de la déviance à un niveau de confiance de confiance de 99%. À ce niveau, la statisque à comparer avec les déviances résiduelles du modèle est de 6.634897. Voici dans le tableau \_ l’analyse pour le modèle complet.

La première variable à enlever est *working.* Le même processus sera exercé jusqu’à ce qu’aucune variable explicative ne soit en-deça de 6.634897. Pour y arriver, nous devrons enlever, dans l’ordre, les variables suivantes : *working, whitec, handdum, addon, public*. Voici le tableau \_ qui présente l’analyse de la variance du modèle courant :

### c) Pearson, déviance et AIC

Le tableau \_ présente les valeurs pour les différentes statistiques nécessaires pour valider le modèle courant.



Tableau \_ : Statistiques de validation d'un modèle

Le tableau \_ présente le sommaire du modèle courant.

Ici, puisque le paramètre de dispersion est de 1, on s’attend à une déviance à peu près égale à n- p' si le modèle est adéquat. Or, à la lumière des résultats présentés ci-haut, le modèle n’est pas adéquat. La déviance résiduelle est de 5009.3 alors que le nombre de degré de liberté résiduel est de 1062.

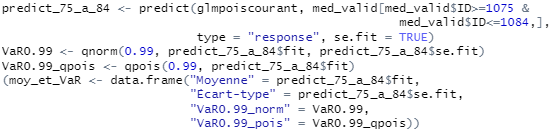
La statistique de Pearson est, quant à elle, tout aussi éloignée du nombre de degré de liberté du modèle.

La statistique AIC est de 7030.4

En conséquence des précédentes dispositions, nous nous rapprochons d’une conclusion quasi-irréversible de rejet de ce modèle pour expliquer la variable réponse de l’étude.

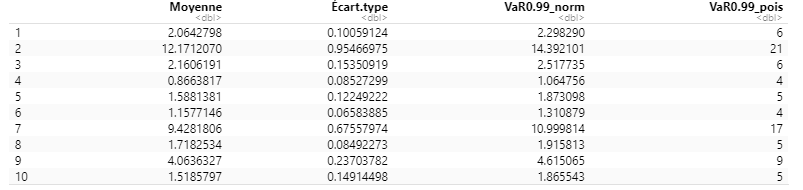
### d) Estimations ponctuelles sur ID 1075 à 1084

Voici le code pour le calcul des paramètres demandés :



Nous avons testé 2 façons de calculer la VaR0.99 pour les raisons suivantes :

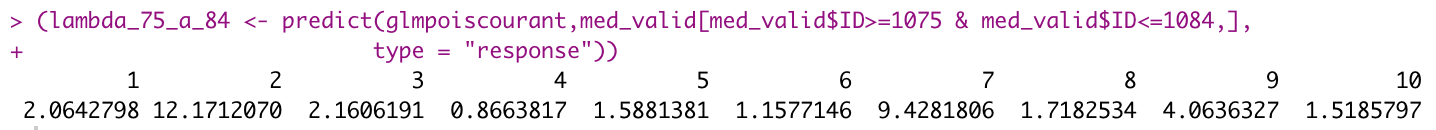
Qnorm : Nous disposons d’une moyenne qui, pour chaque entrée, n’égale pas sa variance



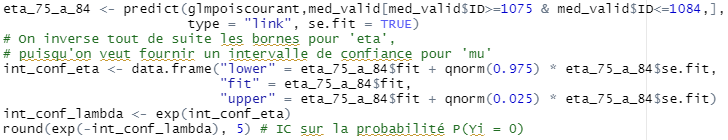
Il est clair que, sachant qu’il s’agit d’un modèle de Poisson, la grande différence qui existe entre la plupart des moyennes et des variances, tel que démontré dans le tableau \_, ne pointe pas dans la direction d’un modèle adapté. Quoi qu’il en soit, les estimations de moyennes pourraient très bien reflétées la réalité.

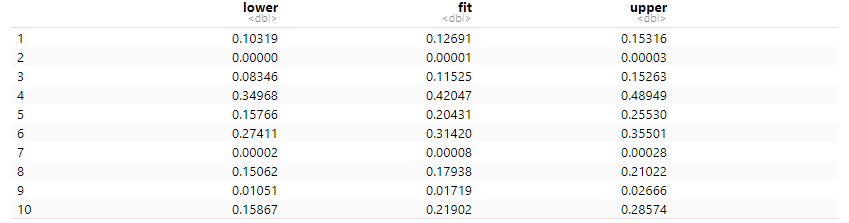
### e) Estimation ponctuelle et intervalle de confiance 95% P(Yi =0)

D’abord, on commence par créer le vecteur des estimations de moyennes pour chacun des assurés.

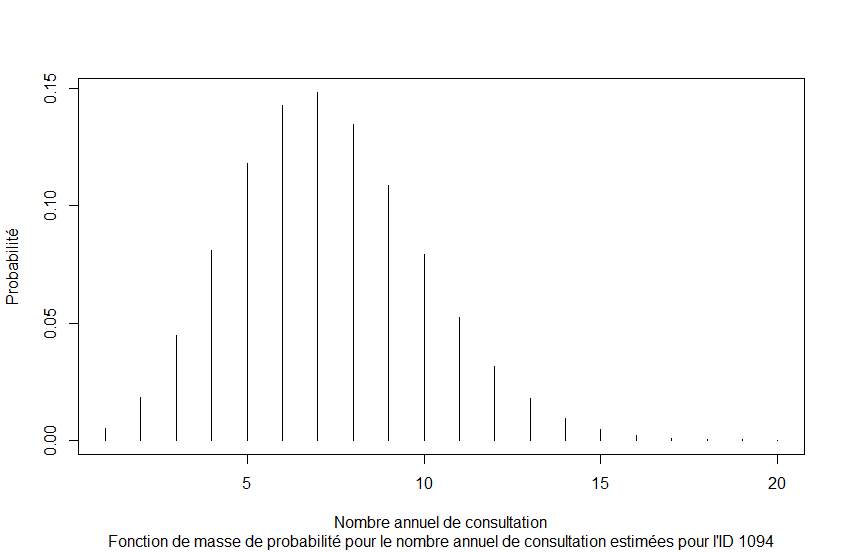
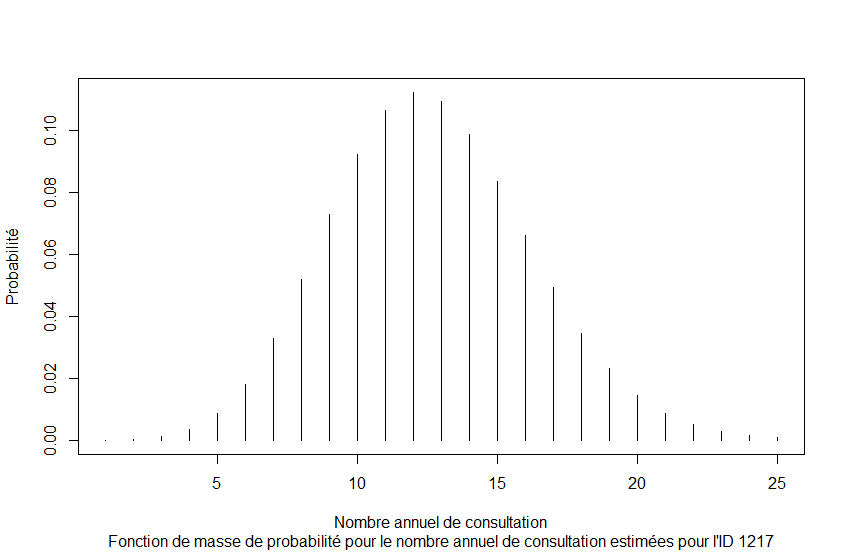


À ce point, puisqu’on nous demande de calculer la probabilité d’aucune visite pour chaque assuré, nous n’avons pas à annualiser la moyenne. Nous procédons directement au calcul sur l’échelle du prédicteur linéaire.

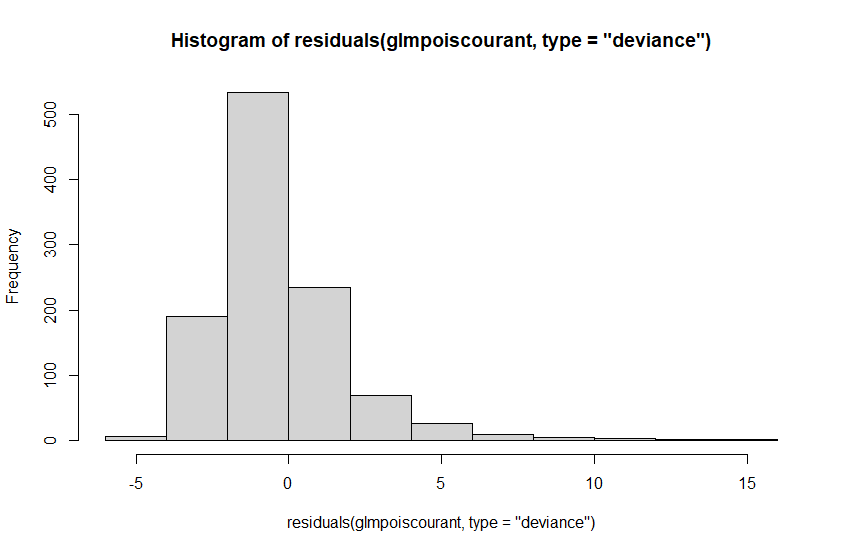




### f) Graphiques des fonctions de masse de probabilité



### g) Graphiques des résidus de Pearson, d’Anscombe et de déviance



## 2. Loi binomial négative avec lien logarithmique

### a) EMV de β et estimation des écart-type

### b) Élaboration du modèle courant

### c) Comparaison des résultats avec le modèle Poisson

### d) Pearson et AIC

### e) Estimation ponctuelle sur ID 1075 à 1084

### f) Estimation ponctuelle et intervalle de confiance 95% P(Yi =0)

### g) Test statistique

## 3. Modèle recommandé

### a) Présentation du modèle

### b) Estimation des paramètres et des écart-types

### c) Statistiques du modèle

### d) Graphiques des fonctions de masse de probabilité

### e) Caractéristique de risque pour VaR0.99

# Second projet : Taux d’abandon au renouvellement de polices d’assurance automobile

### Faits saillants

### 1. Modèle binomial

### a) Expression de probabilité d’abandon

### b) Fonction de lien

### c) Estimation des paramètres et des écart-types

### d) Transformation et traitement des variables explicatives et des interactions

### e) Sélection des variables par la déviance

### f) Courbe ROC

### g) Maximisation de la sensivité et de la spécifité

## 2. Prévisions

### a) Estimation ponctuelle

### b) Interprétation et explication des résultats

### c) Tableau de mauvaise classification

### d) Prévisions et explications