Remarque - Test de surdispersion:

H_0: Pas de surdispersion ($\Rightarrow$ la variance de Y est égale à l'espérance de Y)

H_1: Il y a de la surdispersion

$$\left(\begin{aligned}\Rightarrow Var(Y) &= \mathbb{E}\left(Var(Y|\Theta)\right) + Var\left(\mathbb{E}(Y|\Theta)\right)\\ &= \left(\mathbb{E}(\lambda\times\Theta) + Var(\lambda\times\Theta)\right)\\ &= \lambda\times 1 + \lambda^2\times 1/\alpha\\ &= \lambda\times\left(1+\frac{\lambda}{\alpha}\right)\end{aligned}\right)$$

Procédure:

(1) Obtenir $\ell_{Poisson}(\underline{\beta})$ avec une régression Poisson

(2) Obtenir $\ell_{NB}(\underline{\beta})$ avec une régression Binomiale négative

(3) Calculer $\chi^2_{OBS} = 2\times\left(\ell_{NB}(\underline{\beta}) - \ell_{Poisson}(\underline{\beta})\right)$

(4) On rejette H_0 au niveau de confiance $100\times(1-\alpha)\%$ si

$$\chi^2_{OBS} \geq \chi^2_{1-2\times\alpha}(1)$$

↑ car, le test est "unilatéral" et non bilatéral

B) Loi Gamma :

- Permet de modéliser des cas où Y est continue, mais non-négative (ex : sévérité des sinistres)

- Fonction de log-vraisemblance :

$$f_{Y_i|\underset{\sim}{X}_i}(y_i) = \frac{(\alpha/\mu_i)^{\alpha} \times y_i^{\alpha-1} e^{-(\alpha/\mu_i)\, y_i}}{\Gamma(\alpha)}$$

; où

- $\ln(\mu_i) = \underset{\sim}{X}_i\underset{\sim}{\beta} \overset{\text{lien log}}{\Rightarrow} \mu_i = e^{\underset{\sim}{X}_i\underset{\sim}{\beta}}$
- $\dfrac{1}{\mu_i} = \underset{\sim}{X}_i\underset{\sim}{\beta} \overset{\text{lien inverse}}{\Rightarrow} \mu_i = \dfrac{1}{\underset{\sim}{X}_i\underset{\sim}{\beta}}$.

$$\Rightarrow \mathbb{E}(Y_i|\underset{\sim}{X}_i) = \mu_i = g^{-1}(\underset{\sim}{X}_i\underset{\sim}{\beta})$$

$$\Rightarrow \operatorname{Var}(Y_i|\underset{\sim}{X}_i) = \frac{\mu_i^2}{\alpha}$$

- $\ell(\underset{\sim}{\beta}) = \sum_{i=1}^{n} \ln\left(f_{Y_i|\underset{\sim}{X}_i}(y_i)\right)$

- $S(\underset{\sim}{\beta}) = \dfrac{\partial}{\partial \underset{\sim}{\beta}} \ell(\underset{\sim}{\beta})$

- $I(\underset{\sim}{\beta}) = -\dfrac{\partial^2}{\partial \underset{\sim}{\beta}^2} \ell(\underset{\sim}{\beta})$

- $\hat{\underset{\sim}{\beta}}^{(i+1)} = \hat{\underset{\sim}{\beta}}^{(i)} + I^{-1}(\hat{\underset{\sim}{\beta}}^{(i)})\, S(\hat{\underset{\sim}{\beta}}^{(i)})$.

4.8.2) Algorithme du "scoring":

Est un algorithme itératif pour estimer les paramètres qui peut parfois être une alternative à l'algorithme de Newton-Raphson.

Pour obtenir l'algorithme du scoring, il faut simplement remplacer $I(\hat{\beta}^{(i)})$ par $F(\hat{\beta}^{(i)})$ dans la formule:

$$\tilde{\beta}^{(i+1)} = \hat{\beta}^{(i)} + I^{-1}(\hat{\beta}^{(i)})\, S(\hat{\beta}^{(i)})$$

où

$$F(\hat{\beta}^{(i)}) = E_Y\left(I(\hat{\beta}^{(i)}) \right)$$

L'algorithme du scoring s'écrit donc comme:

$$\hat{\beta}^{(i+1)} = \hat{\beta}^{(i)} + F^{-1}(\hat{\beta}^{(i)})\, S(\hat{\beta}^{(i)})$$

Remarque: En pratique, il est d'usage d'utiliser cet algorithme pour alléger les formules lorsqu'il reste des termes en "y_i" dans la matrice $I(\hat{\beta}^{(i)})$

NOTE: Dans les cas normal, logistique et Poisson, il n'y a pas de "y_i" dans $I(\hat{\beta}^{(i)})$, mais ce n'est pas le cas avec une loi gamma par exemple, ou bien avec la loi géométrique.

exemple : Loi géométrique

(1) $Pr(Y_i = y_i \mid \underset{\sim}{X}_i) = p_i^{y_i - 1} \times (1-p_i)$; $y_i \in \{1, 2, \ldots\}$

(2) Fonction de lien $g(x) = \ln(x-1)$

$$\Rightarrow \quad \ln\left(\mathbb{E}(Y_i \mid \underset{\sim}{X}_i) - 1\right) = \underset{\sim}{X}_i \beta$$

$$\Rightarrow \quad \mathbb{E}(Y_i \mid \underset{\sim}{X}_i) - 1 = e^{\underset{\sim}{X}_i \beta}$$

$$\Rightarrow \quad \mathbb{E}(Y_i \mid \underset{\sim}{X}_i) = 1 + e^{\underset{\sim}{X}_i \beta}$$

$$\Rightarrow \quad \frac{1}{1-p_i} = 1 + e^{\underset{\sim}{X}_i \beta}$$

$$\Rightarrow \quad 1 - p_i = \frac{1}{1+e^{X_i \beta}}$$

$$\Rightarrow \quad p_i = \frac{e^{X_i \beta}}{1+e^{X_i \beta}}$$

<u>Fonction de log-vraisemblance:</u>

$$\ell(\underset{\sim}{\beta}) = \sum_{i=1}^{n} \ln\left(Pr(Y_i = y_i \mid \underset{\sim}{X}_i)\right)$$

$$= \sum_{i=1}^{n} \left\{ (y_i - 1)\ln(p_i) + \ln(1-p_i) \right\}$$

$$= \sum_{i=1}^{n} \left\{ (y_i - 1) \ln\left(\frac{e^{\underset{\sim}{X}_i \underset{\sim}{\beta}}}{1+e^{\underset{\sim}{X}_i \underset{\sim}{\beta}}} \right) + \ln\left(\frac{1}{1+e^{\underset{\sim}{X}_i \underset{\sim}{\beta}}} \right) \right\}$$

$$= \sum_{i=1}^{n} \left\{ (y_i - 1) \left[\underset{\sim}{X}_i \underset{\sim}{\beta} - \ln\left(1+e^{\underset{\sim}{X}_i \underset{\sim}{\beta}}\right) \right] - \ln\left(1+e^{\underset{\sim}{X}_i \underset{\sim}{\beta}}\right) \right\}$$

$$= \sum_{i=1}^{n} \left\{ (y_i - 1)(\underset{\sim}{X}_i \underset{\sim}{\beta}) - y_i \times \ln\left(1+e^{\underset{\sim}{X}_i \underset{\sim}{\beta}}\right) \right\}$$

Vecteur score:

$$[S(\underset{\sim}{\beta})]_j = \frac{\partial}{\partial \beta_j} \ell(\underset{\sim}{\beta}) = \sum_{i=1}^{n} \left\{ (y_i - 1) X_{ij} - y_i \times X_{ij} \frac{e^{\underset{\sim}{X}_i \underset{\sim}{\beta}}}{1+e^{\underset{\sim}{X}_i \underset{\sim}{\beta}}} \right\}$$

$$= \ldots$$

$$= \sum_{i=1}^{n} \left\{ \frac{y_i}{1+e^{\underset{\sim}{X}_i \underset{\sim}{\beta}}} - 1 \right\} \times X_{ij}$$

Matrice d'information de Fisher:

$$[I(\underset{\sim}{\beta})]_{j,k} = \frac{\partial^2 \ell(\underset{\sim}{\beta})}{\partial \beta_j \partial \beta_k} = \frac{\partial}{\partial \beta_k} \left(\frac{\partial}{\partial \beta_j} \ell(\underset{\sim}{\beta}) \right) = \frac{\partial}{\partial \beta_k} [S(\underset{\sim}{\beta})]_j$$

$$= \sum_{i=1}^{n} \left\{ y_i \times \left(\frac{e^{\underset{\sim}{X}_i \underset{\sim}{\beta}}}{1+e^{\underset{\sim}{X}_i \underset{\sim}{\beta}}} \right) \times \left(\frac{1}{1+e^{\underset{\sim}{X}_i \underset{\sim}{\beta}}} \right) \times X_{ij} \times X_{ik} \right\}$$

... il y a des "y_i" dans $I(\underset{\sim}{\beta})$!

Matrice F :

$$[F(\beta)]_{j,k} = \mathbb{E}_Y\left([I(\beta)]_{j,k}\right)$$

$$= \mathbb{E}_Y\left(\sum_{i=1}^{n}\left\{y_i \times \left(\frac{e^{X_i\beta}}{1+e^{X_i\beta}}\right) \times \left(\frac{1}{1+e^{X_i\beta}}\right) \times X_{ij} \times X_{ik}\right\}\right)$$

$$= \sum_{i=1}^{n}\left\{ \mathbb{E}(Y_i \mid X_i) \times \left(\frac{e^{X_i\beta}}{1+e^{X_i\beta}}\right) \times \left(\frac{1}{1+e^{X_i\beta}}\right) X_{ij} X_{ik}\right\}$$

$$= \sum_{i=1}^{n}\left\{\left(\frac{1}{1-p_i}\right) \times \left(\frac{e^{X_i\beta}}{1+e^{X_i\beta}}\right) \times \left(\frac{1}{1+e^{X_i\beta}}\right) X_{ij} \times X_{ik}\right\}$$

$$= \sum_{i=1}^{n}\left\{\left(1+e^{X_i\beta}\right)\left(\frac{e^{X_i\beta}}{1+e^{X_i\beta}}\right)\left(\frac{1}{1+e^{X_i\beta}}\right) X_{ij} \times X_{ik}\right\}$$

$$= \sum_{i=1}^{n}\left(\frac{e^{X_i\beta}}{1+e^{X_i\beta}}\right) \times X_{ij} \times X_{ik} = \sum_{i=1}^{n} X_{ij} \times \omega_i \times X_{ik}$$

... plus simple, car non fonction des "y_i".

... $F(\beta)$ peut donc s'écrire comme suit :

$$F(\beta) = X' W X \; ; \; W = \begin{bmatrix} \omega_1 & 0 & \cdots & 0 \\ 0 & \omega_2 & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & \omega_n \end{bmatrix}$$

4.8.3) Variable "offset" :

... est une variable explicative dont le "β" est 1.

Utilisation en actuariat :

Modélisation de la fréquence des sinistres à partir de la loi Poisson

exemple :

i	U_i	N_i	X_{i1}	X_{i2}	...	X_{ip}
1	0.5	1				
2	1.0	2				
3	1.0	0				
4	0.25	1				
5	0.05	0				

- U_i : % de l'année où l'assuré i a détenu une police en vigueur

  ≈ les unités d'exposition.

- N_i : Nombre de sinistre de l'assuré i.

Modèle GLM :

$$(N_i \mid U_i, \underline{X}_i) \sim \text{Poisson}(\lambda_i)$$

$$\lambda_i = U_i \times e^{X_i \beta}$$

Idée sous-jacente à ce modèle:

Un assuré qui a détenu une police en vigueur pendant quelque jours a considérablement moins de chances d'avoir des sinistres qu'un assuré qui a conservé sa police en vigueur pendant toute l'année!

Interprétation:

- $\lambda_i = U_i \times e^{X_i \beta}$: Nombre de sinistre espéré pour l'assuré i.

- $\dfrac{\lambda_i}{U_i} = e^{X_i \beta}$: Fréquence de sinistre espéré pour l'assuré i

Variable "offset":

$$\begin{aligned}
\lambda_i &= U_i \times e^{X_i \beta} \\
&= e^{\ln(U_i)} \times e^{\beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip}} \\
&= e^{\ln(U_i) + \beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip}}
\end{aligned}$$

... en posant $O_i = \ln(U_i)$; alors on a que:

$$\boxed{\lambda_i = e^{1 \times O_i + \beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip}}}$$

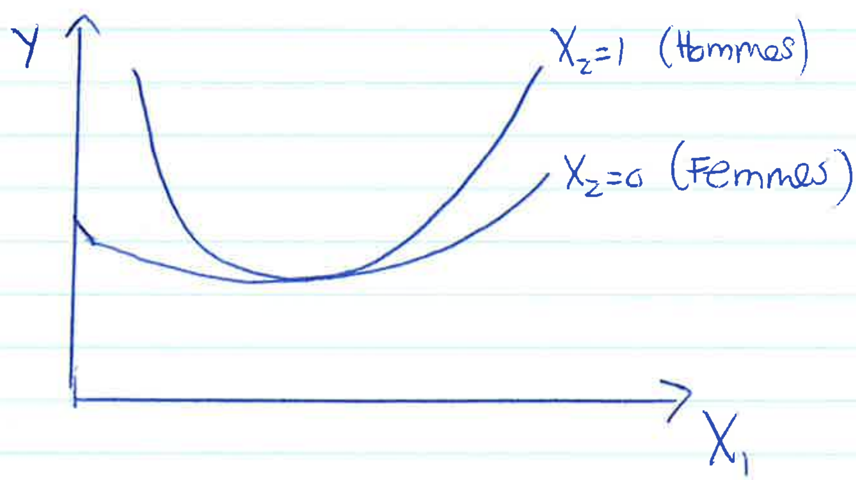
Ici, O_i est appelé "variable offset" !

4.8.4) Ajout d'une interaction entre 2 variables explicatives:

... si on pense que l'effet combiné de 2 variables explicatives doit être différent de l'addition des effets individuels de ces 2 variables

exemple: Y: Coûts d'assurance auto
X_1: Age de l'assuré au carré
X_2: Indicateur que l'assuré est un homme

En pratique, on observe souvent la tendance suivante:

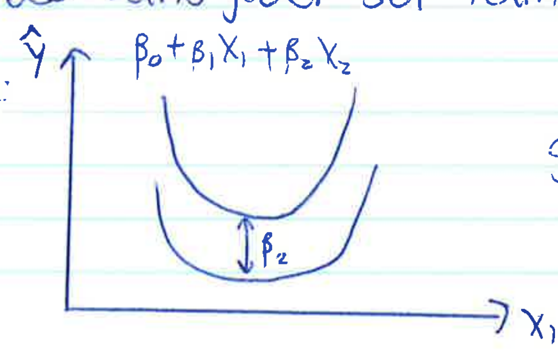


Or, il est impossible d'obtenir le graphique précédent avec une structure linéaire de la forme:

$$\beta_0 + \beta_1 \cdot X_1 + \beta_2 \cdot X_2$$

Car le β_2 ne fait qu'induire un effet de translation sur les courbes sans jouer sur l'amplitude:

ex:



$$g(E(Y|\underset{\sim}{X_i})) = \begin{cases} \beta_0+\beta_1 X_1 \text{, si femme} \\ (\beta_0+\beta_2)+\beta_1 X_1 \text{, si homme} \end{cases}$$

Pour permettre à l'amplitude de varier selon le sexe, il faut ajouter une interaction à la structure linéaire précédente, soit :

$$\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \underbrace{\beta_3 \times X_1 \times X_2}_{\text{Partie "interaction" entre } X_1 \text{ et } X_2}$$

Ainsi:

$$g(\mathbb{E}(Y_i | \underset{\sim}{X}_i)) = \begin{cases} \beta_0 + \beta_1 X_1 & \text{, si femme} \\ (\beta_0 + \beta_2) + (\beta_1 + \beta_3) X_1 & \text{, si homme} \end{cases}$$

Remarque importante:

Tout comme le traitement des variables non-numériques, l'ajout d'une interaction ne nécessite que de modifier la matrice shéma X et le vecteur $\underset{\sim}{\beta}$:

Sans interaction:

$$X = \begin{array}{cccc} & X_1 & X_2 \\ \left[\begin{array}{c} 1 \\ 1 \\ 1 \\ \vdots \end{array}\right. & \begin{array}{c} 60 \\ 30 \\ 20 \\ \vdots \end{array} & \left.\begin{array}{c} 0 \\ 1 \\ 0 \\ \vdots \end{array}\right] \end{array} ; \quad \underset{\sim}{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}$$

Avec interaction:

$$X = \begin{array}{cccc} & X_1 & X_2 & X_3 = X_1 \times X_2 \\ \left[\begin{array}{c} 1 \\ 1 \\ 1 \\ \vdots \end{array}\right. & \begin{array}{c} 60 \\ 30 \\ 20 \\ \vdots \end{array} & \begin{array}{c} 0 \\ 1 \\ 0 \\ \vdots \end{array} & \left.\begin{array}{c} 0 \\ 30 \\ 0 \\ \vdots \end{array}\right] \end{array} ; \quad \underset{\sim}{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}$$

4.8.5) Analyse des résidus:

Toute la théorie introduite au chapitre 3 reste valide pour les GLM à l'exception du fait qu'il soit nécessaire de définir un résidu plus général.

Résidus de déviance:

$$\hat{\varepsilon}_D = \text{signe}(y_i - \hat{\mu}_i) \times \sqrt{2 \times \left(\ell(\underset{\sim}{y}, \underset{\sim}{y})_i - \ell(\underset{\sim}{y}, \underset{\sim}{\mu})_i \right)}$$

...les i^e termes dans la somme de $\ell(\underset{\sim}{\beta})$