MODELISATION DE DONNÉES QUALITATIVES

REGRESSION LOGISTIQUE SIMPLE

Pierre-Louis Gonzalez

MODELES A REPONSE DICHOTOMIQUE

Quelques applications:
Y est dichotomique: succès ou échec, présence ou absence.

- **Un organisme de crédit doit–il accorder un prêt à l'un de ses clients ?**
- **Une entreprise présente t-elle des risques de faillite à moyen terme (2ans) ?**
- **Un patient est-il atteint d'une certaine maladie ?**
- **Doit-on déclencher une alerte à la pollution atmosphérique ?**

LA RÉGRESSION LOGISTISQUE SIMPLE
(cas Y binaire)

I LE MODÈLE LOGISTIQUE

1 - Les données

Y = variable à expliquer binaire

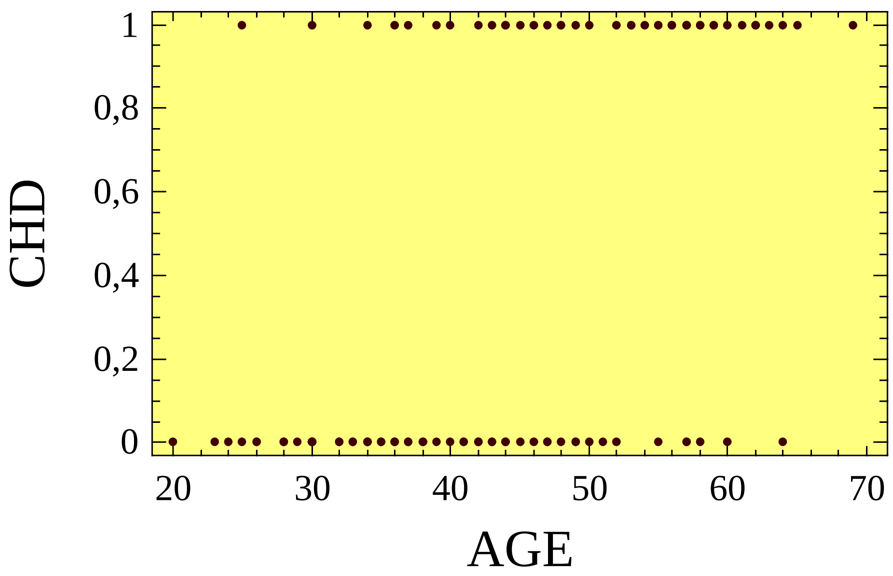
(Ex : présence/absence d'une maladie cardiaque)

X = variable explicative quantitative

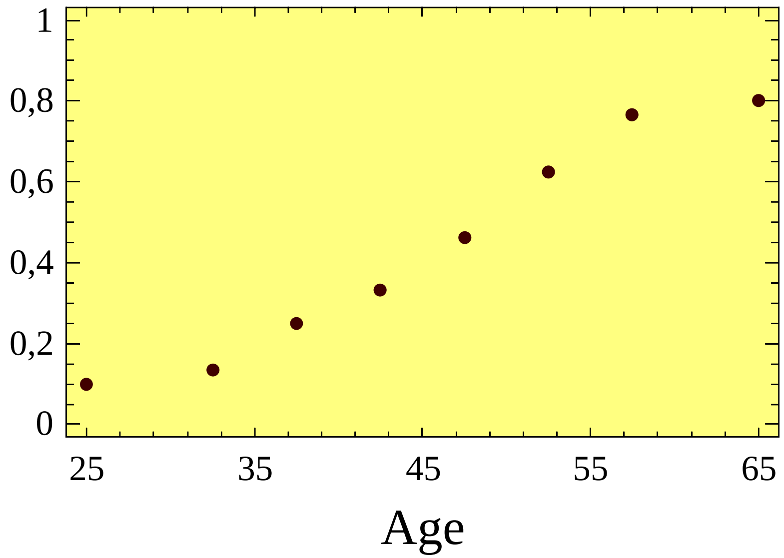
(Ex : âge)

II.1 Etude des données

- Le graphique 4 ne montre pas clairement l’existence d’une liaison entre Y et X.



- Par contre si l'on utilise la variable âge découpée en classes et la proportion de malades par classe, la liaison entre Y et X apparaît plus clairement sous la forme d'une courbe en S.



2 - Le modèle logistique

Objectif Modéliser $\Pi(\mathbf{x}) = \text{Prob}(\mathbf{Y} = \mathbf{1} / \mathbf{X} = \mathbf{x})$

2.1 Modèle linéaire

$$\Pi(\mathbf{x}) = \beta_0 + \beta_1 \mathbf{x}$$

Ce modèle convient mal pour deux raisons :

- $\Pi(\mathbf{x})$ sort de $[\mathbf{0,1}]$
- $\Pi'(\mathbf{x})$ doit tendre vers **0** lorsque $\Pi(\mathbf{x})$ tend vers **0** ou **1**

  $\Pi(\mathbf{x})$ est une courbe en **S**

2.2 Modèle logistique

$$\Pi(\mathbf{x}) = \frac{\mathbf{e}^{\beta_0 + \beta_1 \mathbf{x}}}{\mathbf{1} + \mathbf{e}^{\beta_0 + \beta_1 \mathbf{x}}}$$

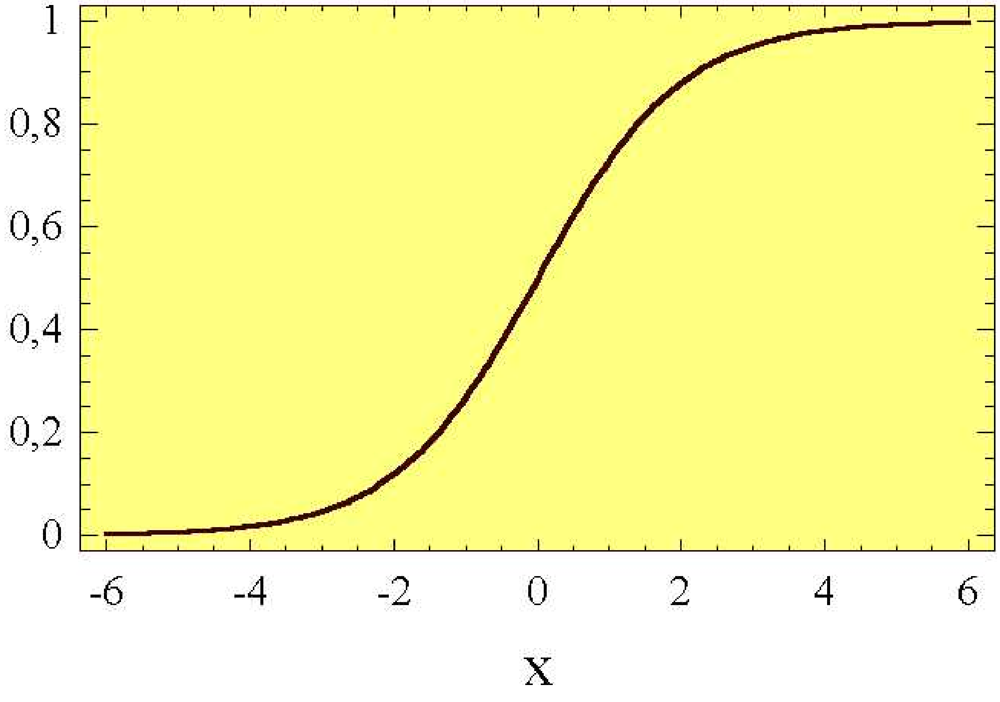
$$\Pi'(\mathbf{x}) = \beta_1 \, \Pi(\mathbf{x}) \left(1 - \Pi(x)\right)$$

Donc $\Pi'(\mathbf{x}) \rightarrow \mathbf{0}$ lorsque $\Pi(\mathbf{x}) \rightarrow \mathbf{0}$ ou $\mathbf{1}$

Justifications concernant le choix de la fonction logistique

La fonction logistique est définie par :

$$F(x) = 1 / (1 + e^{-x})$$



- Cette fonction est bien adaptée à la modélisation de probabilités car elle prend ses valeurs entre 0 et 1 **selon une courbe en S**.

  Son utilisation est par exemple indiquée lors de la modélisation du risque individuel de développer une maladie dans les études épidémiologiques.

- En effet, en considérant que la variable x représente un indice résultant de la combinaison de plusieurs facteurs de risque, on peut interpréter F(x) comme le risque d'être atteint de cette maladie.

  Dans ce contexte le risque est minimal pour de faibles valeurs de x.
  Il augmente pour les valeurs intermédiaires de x.
  Il apparaît proche de 1 pour des valeurs plus élevées de x.

3 - Estimation des paramètres du modèle logistique

Données

X	Y
x_1	y_1
x_2	y_2
$\vdots$	$\vdots$
x_i	y_i
$\vdots$	$\vdots$
x_n	y_n

$$\Pi(x_i) = \Pi_i = P(Y = 1 | X = x_i)$$

$$= \frac{e^{\beta_0 + \beta_1 x_i}}{1 + e^{\beta_0 + \beta_1 x_i}}$$

Vraisemblance

$$\mathbf{L}(\mathbf{B}) = \prod_{i=1}^{n} \Pi(\mathbf{x}_i)^{y_i} \left(1 - \Pi(\mathbf{x}_i)\right)^{1-y_i}$$

$\Big($Ceci résulte du fait que :

$$\mathbf{f}(\mathbf{y}_i, \beta) = \text{Prob}\left[\mathbf{Y} = \mathbf{y}_i \,\middle|\, \mathbf{X} = \mathbf{x}_i, \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} \right]$$

$$= \begin{cases} \Pi(\mathbf{x}_i) \text{ pour } \mathbf{y}_i = 1 \\ 1 - \Pi(\mathbf{x}_i) \text{ pour } \mathbf{y}_i = 0 \end{cases} \Big)$$

Log de la vraisemblance

$$\mathrm{Log}\, \mathbf{L}(\beta) = \sum_{i=1}^{n} \mathbf{y}_i \; \mathrm{Log}\, \Pi_i(\mathbf{x}) + (\mathbf{1} - \mathbf{y}_i) \; \mathbf{Log}(\mathbf{1} - \Pi_i(\mathbf{x}))$$

Maximum de vraisemblance

On obtient $\hat{\beta}$ en annulant $\dfrac{\partial \; \mathrm{Log}\, \mathbf{L}(\beta)}{\partial \beta}$

$$\begin{cases} \dfrac{\partial \; \mathrm{Log}\, \mathbf{L}(\beta)}{\partial \beta_0} = \displaystyle\sum_{i=1}^{n} (\mathbf{y}_i - \Pi(\mathbf{x}_i)) \\ \qquad\qquad = \displaystyle\sum_{i=1}^{n} \left(\mathbf{y}_i - \dfrac{\mathbf{e}^{\beta_0 + \beta_1 \mathbf{x}_i}}{\mathbf{1} + \mathbf{e}^{\beta_0 + \beta_1 \mathbf{x}_i}} \right) = \mathbf{0} \\ \dfrac{\partial \; \mathrm{Log}\, \mathbf{L}(\beta)}{\partial \beta_1} = \displaystyle\sum_{i=1}^{n} \mathbf{x}_i (\mathbf{y}_i - \Pi(\mathbf{x}_i)) \\ \qquad\qquad = \displaystyle\sum_{i=1}^{n} \mathbf{x}_i \left(\mathbf{y}_i - \dfrac{\mathbf{e}^{\beta_0 + \beta_1 \mathbf{x}_i}}{\mathbf{1} + \mathbf{e}^{\beta_0 + \beta_1 \mathbf{x}_i}} \right) = \mathbf{0} \end{cases}$$

Résultats

$$\hat{\beta}_0 = -5{,}3095$$

$$\hat{\beta}_1 = 0{,}1109$$

$$-\,2\ \mathrm{Log}\ L = 107{,}35 \quad \text{est minimum}$$

Modèle estimé : $$\hat{\Pi}(x) = \frac{e^{-5{,}3095+0{,}1109x}}{1+e^{-5{,}3095+0{,}1109x}}$$

La probabilité d’être atteint de la maladie augmente avec l’âge.

4 - Calcul des écarts-types des estimateurs des paramètres

En notant $\hat{\Pi}_i = \hat{\Pi}(\mathbf{x}_i)$, on a :

$$\hat{V}(\hat{\beta}) = \left[\frac{-\partial^2 \text{ Log } \mathbf{L}(\beta)}{\partial \beta^2}\right]^{-1}_{\beta=\hat{\beta}} = \begin{bmatrix} \sum_{i=1}^{n} \hat{\Pi}_i(1-\hat{\Pi}_i) & \sum_{i=1}^{n} x_i \hat{\Pi}_i(1-\hat{\Pi}_i) \\ \sum_{i=1}^{n} x_i \hat{\Pi}_i(1-\hat{\Pi}_i) & \sum_{i=1}^{n} x_i^2 \hat{\Pi}_i(1-\hat{\Pi}_i) \end{bmatrix}^{-1}$$

$$= \left(\begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}' \begin{bmatrix} \hat{\Pi}_1(1-\hat{\Pi}_1) & & 0 \\ & \ddots & \\ 0 & & \hat{\Pi}_n(1-\hat{\Pi}_n) \end{bmatrix} \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \right)^{-1}$$

$$= (X'VX)^{-1}$$

Résultats

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	-5.3095	1.1337	21.9350	0.0001	.	.
AGE	1	0.1109	0.0241	21.2541	0.0001	0.716806	1.117

$$\hat{\mathbf{V}}\left(\hat{\beta}\right) = \begin{bmatrix} \mathbf{1{,}285} & \mathbf{-0{,}0266} \\ \mathbf{-0{,}0266} & \mathbf{0{,}0005388} \end{bmatrix}$$

D'où

$$\mathbf{s}\left(\hat{\beta}_0\right) = \sqrt{\mathbf{1{,}285}} = \mathbf{1{,}1337}$$

$$\mathbf{s}\left(\hat{\beta}_1\right) = \sqrt{\mathbf{0{,}0005388}} = \mathbf{0{,}0241}$$

5 - Utilisations des écarts-types

5.1 Test de nullité (ou de signification) d'un coefficient β_j au niveau α

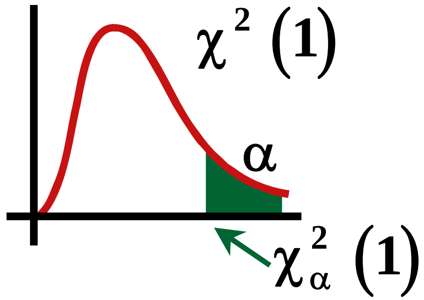
$$\begin{cases} H_0 : \beta_j = 0 \\ H_1 : \beta_j \neq 0 \end{cases}$$

Test de Wald :

$$\frac{\hat{\beta}_j^2}{s^2\left(\hat{\beta}_j\right)} \longrightarrow \chi^2(1) \qquad \text{sous } \mathbf{H_0}$$

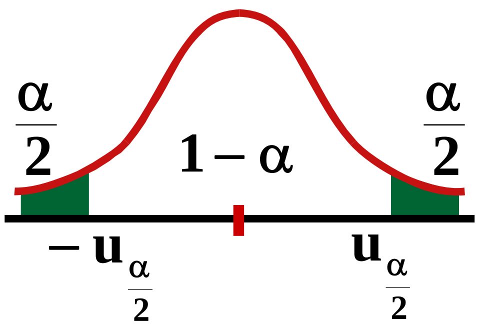
On rejette $\mathbf{H_0}$ si :

$$W_j = \frac{\hat{\beta}_j^2}{s^2\left(\hat{\beta}_j\right)} > \chi^2_\alpha(1)$$



$$\Leftrightarrow \quad \mathbf{Pr}\underbrace{\left[\chi^2_{(1)} > \mathbf{w}_j\right]}_{\text{P-value}} < \alpha$$

$$\Leftrightarrow \quad \frac{\left|\hat{\beta}_j\right|}{\mathbf{s}\left(\hat{\beta}_j\right)} > \mathbf{u}_{\frac{\alpha}{2}}$$



5.2 Intervalle de confiance de $\Pi(x)$

$$\Pi(x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}}$$

On déduit un intervalle de confiance de $\Pi(x)$ de l'intervalle de confiance de $g(x) = \beta_0 + \beta_1 x.$

I.C. à 95 % de $g(x) = \beta_0 + \beta_1 x.$

$$\hat{\beta}_0 + \hat{\beta}_1(x) \pm 1{,}96\, \hat{\sigma}\left(\hat{\beta}_0 + \hat{\beta}_1 x\right)$$

avec $$\hat{\sigma}\left(\hat{\beta}_0 + \hat{\beta}_1 x\right) = (1, x)\ \hat{V}\left(\hat{\beta}\right)\begin{pmatrix}1\\ x\end{pmatrix}$$

Exemple

$$x = 20 \text{ ans}$$

$$\hat{g}(20) = \hat{\beta}_0 + \hat{\beta}_1\ 20 = -5{,}3095 + 0{,}1109 * 20$$

$$= \underline{-3{,}0915}$$

$$\hat{\Pi}(20) = \frac{e^{-3{,}0915}}{1 + e^{-3{,}0915}} = \frac{0{,}0454338}{1{,}0454338} = \underline{0{,}0434592}$$

$$\hat{\sigma}^2\left(\hat{\beta}_0 + \hat{\beta}_1 20\right) = (1{,}20) \begin{bmatrix} 1{,}2852 & -\ 0{,}0267 \\ -\ 0{,}0267 & 0{,}0005788 \end{bmatrix} \begin{pmatrix} 1 \\ 20 \end{pmatrix}$$

$$= \underline{0{,}448}$$

Intervalle de confiance à 95 % de $g(20) = \beta_0 + \beta_1 20$

$$-3{,}0915 \;\pm\; 1{,}96\,\sqrt{0{,}448} \;=\; [-4{,}4034\;;\;-1{,}7796]$$

Intervalle de confiance à 95 % de $\Pi(20) = \text{Prob}\,(Y = 1 | X = 1)$

$$\text{Inf}\,(95\ \%) = \frac{e^{-4{,}4034}}{1 + e^{-4{,}4034}} = 0{,}012$$

$$\text{Sup}\,(95\ \%) = \frac{e^{-1{,}7796}}{1 + e^{-1{,}7796}} = 0{,}144$$

6 - Autres tests

6.1 Test du rapport de vraisemblance

Modèle $$\Pi(x_i) = \frac{e^{\beta_0+\beta_1 x_i}}{1+e^{\beta_0+\beta_1 x_i}}$$

Test de l'influence de X sur Y

$$\begin{cases} H_0 : \beta_1 = 0 \\ H_1 : \beta_1 \neq 0 \end{cases}$$

Statistique utilisée

$$\Lambda = -2 \text{ Log}\left[\frac{L\left(\tilde{\beta}, \beta_1 = 0\right)}{L\left(\hat{\beta}_0, \hat{\beta}_1\right)}\right]$$

Calcul du Log (vraisemblance) sous H_0

$$\text{Sous } \mathbf{H_0} : \Pi(\mathbf{x_i}) = \frac{\mathbf{e}^{\beta_0+\beta_1 x_i}}{1+\mathbf{e}^{\beta_0+\beta_1 x_i}} = \frac{\mathbf{e}^{\beta_0}}{1+\mathbf{e}^{\beta_0}}$$

$$\hat{\Pi}(\mathbf{x_i}) = \frac{\mathbf{e}^{\tilde{\beta}_0}}{1+\mathbf{e}^{\tilde{\beta}_0}} = \frac{\mathbf{n_1}}{\mathbf{n}}$$

$$\text{où } \mathbf{n_1} = \text{nombre de } (\mathbf{y_i} = 1)$$

$$\mathbf{Log}\left(\mathbf{L}\left(\tilde{\beta}_0, \beta_1 = 0\right)\right) = \sum_{i=1}^{n}\left[\mathbf{y_i}\ \mathbf{Log}\ \hat{\Pi}_i + (1-\mathbf{y_i})\ \mathbf{Log}\left(1-\hat{\Pi}_i\right)\right]$$

$$= \mathbf{n_1}\ \mathbf{Log}\ \frac{\mathbf{n_1}}{\mathbf{n}} + \mathbf{n_0}\ \mathbf{Log}\ \frac{\mathbf{n_0}}{\mathbf{n}}$$

$$\text{où } \mathbf{n_0} = \text{nombre de } (\mathbf{y_i} = 0)$$

Résultat

$$\Lambda = -2\ \mathrm{Log}\left[\frac{L\left(\tilde{\beta}_0, \beta_1 = 0\right)}{L\left(\hat{\beta}_0, \hat{\beta}_1\right)}\right]$$

$$= -2\ \mathbf{Log}\ (\text{vraisemblance sans la variable})$$
$$-\left(-\ 2\ \mathbf{Log}\ (\text{vraisemblance avec la variable})\right)$$

$$\longrightarrow \quad \chi^2(1) \text{ sous } \mathbf{H_0}$$

Exemple

$$\Lambda = 136{,}663 - 107{,}353 = 29{,}30$$

$\Rightarrow$ âge très significatif

6.2 Test du score

Vecteur score :

$$\mathbf{u}(\beta) = \begin{bmatrix} \dfrac{\partial \, \mathbf{Log\ L}(\beta)}{\partial \beta_0} \\[2ex] \dfrac{\partial \, \mathbf{Log\ L}(\beta)}{\partial \beta_1} \end{bmatrix}$$

$$\mathbf{u}(\beta) \;\longrightarrow \mathrm{N}\left(\mathbf{0}, \mathbf{I}(\beta)\right)$$

où $\mathbf{I}(\beta) = \mathbf{E}\left[\dfrac{-\partial^2 \, \mathbf{Log\ L}(\beta)}{\partial \beta^2}\right]$

Test $\begin{cases} \mathbf{H_0 : \beta_1 = 0} \\ \mathbf{H_1 : \beta_1 \neq 0} \end{cases}$

On estime β sous $\mathbf{H_0}$

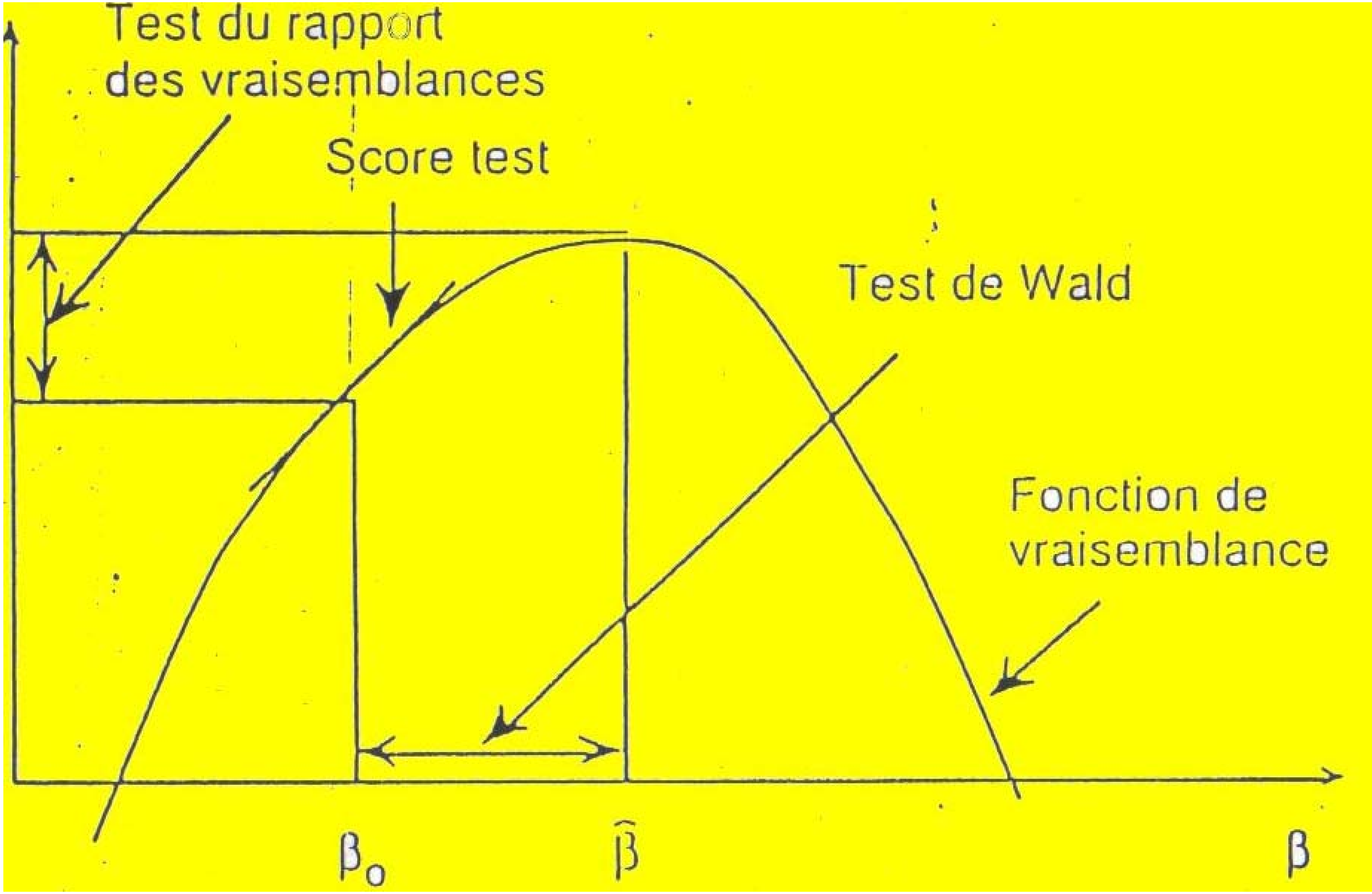
$$\Rightarrow \hat{\beta}_{H_0} = \left(\hat{\beta}_0, 0\right) = \left(\mathbf{Log}\ \frac{\mathbf{n_1}}{\mathbf{n}}, \mathbf{0}\right)$$

Statistique utilisée

$$\text{Score} = U\left(\hat{\beta}_{H_0}\right)' \hat{I}\left(\hat{\beta}_{H_0}\right)^{-1} U\left(\hat{\beta}_{H_0}\right)$$

$$= \frac{\left(\Sigma\, x_i\left(y_i - \bar{y}\right)\right)^2}{\bar{y}\left(1 - \bar{y}\right) \Sigma\left(x_i - \bar{x}\right)^2} \longrightarrow \chi^2(1) \text{ sous } \mathbf{H_0}$$

Exemple Score $= \mathbf{26{,}4} \Rightarrow$ âge très significatif


Test du rapport
des vraisemblances
Score test
Test de Wald
Fonction de
vraisemblance
β0
β̂
β

II - ANALYSE DES RÉSIDUS, DES OBSERVATIONS

1 - Analyse des résidus

Résidu de Pearson

Modèle : $\mathbf{Y_i = \Pi_i + \varepsilon_i}$

$$\Rightarrow \ \boldsymbol{\varepsilon_i = 1 - \Pi_i} \text{ avec une proba } \boldsymbol{\Pi_i}$$

$$\text{et } \boldsymbol{\varepsilon_i = -\Pi_i} \text{ avec une proba } \boldsymbol{1 - \Pi_i}$$

$$\mathbf{E(\varepsilon_i) = (1 - \Pi_i)\Pi_i - \Pi_i(1 - \Pi_i) = 0}$$

$$\mathbf{V(\varepsilon_i) = (1 - \Pi_i)^2\Pi_i + \Pi_i^2(1 - \Pi_i) = \Pi_i(1 - \Pi_i)}$$

On définit le **résidu de Pearson** par :

$$\mathbf{r_i} = \frac{\mathbf{y_i} - \hat{\Pi}_i}{\sqrt{\hat{\Pi}_i\left(1 - \hat{\Pi}_i\right)}} \quad \text{à comparer à 2}$$

Résultat :

$$\chi^2 = \sum_{i=1}^{n} \mathbf{r_i^2} \longrightarrow \chi^2_{n-2} \quad \text{si le modèle étudié est exact.}$$

Déviance

$$\text{Déviance} = \mathbf{-2\ Log}\left(\text{vraisemblance modèle étudié}\right)$$

$$= -2\sum_{i=1}^{n}\left(y_i \ \text{Log}\ \hat{\Pi}_i + (1-y_i)\ \text{Log}\left(1-\hat{\Pi}_i\right)\right)$$

$$\mathbf{D = -2\ Log\ L \approx} \text{ somme des carrés résiduelle}$$

Résidu-déviance :

$$d_i = \begin{cases} \sqrt{2y_i\ \left|\text{Log}\ \hat{\Pi}_i\right|} = \sqrt{2\ \left|\text{Log}\ \Pi_i\right|} & \text{pour } y_i = 1 \\ -\sqrt{2\ (1-y_i)\ \left|\text{Log}\ \left(1-\hat{\Pi}_i\right)\right|} = -\sqrt{2\ \left|\text{Log}\ \left(1-\hat{\Pi}_i\right)\right|} & \text{pour } y_i = 0 \end{cases}$$

Résultat :

$$\mathbf{D} = \sum_{i=1}^{n} d_i^2 = -2\ \text{Log}\ L \longrightarrow \chi^2_{(n-2)} \quad \text{si le modèle étudié est exact.}$$

2 - Analyse des observations

Levier

$$\mathbf{H} = \mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\ \mathbf{X}'$$ en régression multiple

$$\mathbf{H} = \mathbf{V}^{1/2}\mathbf{X}\left(\mathbf{X}'\mathbf{V}\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{V}^{1/2}$$ en régression logistique

où $\mathbf{X} = \begin{pmatrix} \mathbf{1} & \mathbf{x_1} \\ \vdots & \vdots \\ \mathbf{1} & \mathbf{x_n} \end{pmatrix}$ et $\mathbf{V} = \begin{pmatrix} \hat{\Pi}_1\left(1-\hat{\Pi}_1\right) & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \hat{\Pi}_n\left(1-\hat{\Pi}_n\right) \end{pmatrix}$

On définit le **levier** par :

$$\mathbf{h_i} = \hat{\Pi}_i\left(1-\hat{\Pi}_i\right)\left(1,\mathbf{x_i}\right)\left(\mathbf{X}'\mathbf{V}\mathbf{X}\right)^{-1}\begin{pmatrix} \mathbf{1} \\ \mathbf{x_i} \end{pmatrix}$$

Résultats : $\frac{1}{n} \leq h_i \leq 1$ $\qquad$ $\overline{h} = \frac{2}{n}$

Mesure l'éloignement d'une observation par rapport aux autres dans l'espace des variables explicatives lorsque $0{,}1 \leq \hat{\Pi}_i \leq 0{,}9$.

On compare h_i à $2\overline{h} = \frac{4}{n}$.

Influence de chaque observation sur le calcul de $\hat{\beta}$

On note $\hat{\beta}_{(-i)}$ = estimation de β sans utiliser l'observation $\mathbf{i}$

$$\mathbf{C_i} = \left(\hat{\beta} - \hat{\beta}_{(-i)}\right)' \mathbf{X'VX} \left(\hat{\beta} - \hat{\beta}_{(-i)}\right)$$

$$= \frac{\mathbf{r_i^2 h_i}}{\left(\mathbf{1 - h_i}\right)^2}$$

Influence de chaque observation sur le χ^2 de Pearson

$$\Delta_i \chi^2 = \frac{\mathbf{r_i^2}}{\mathbf{1 - h_i}}$$

Influence de chaque observation sur la déviance

$$\Delta_i D = d_i^2 + \frac{r_i^2 h_i}{1 - h_i}$$

Influence de chaque observation sur le calcul de $\hat{\beta}_j$

$$DFBETA_j(-i) = \frac{\hat{\beta}_j - \hat{\beta}_j(-i)}{\hat{\sigma}\left(\hat{\beta}_j\right)}$$

III - REMARQUES

1 Le cas où la variable indépendante est à plusieurs modalités est traité dans le cadre de la **régression logistique multiple**.

En effet, on remplace la colonne de la variable qualitative codée par le tableau des indicatrices de ses modalités.

Pour éviter les **problèmes d'indétermination**, on supprime une des indicatrices des calculs (souvent on supprime celle qui correspond à la situation la plus courante : situation de référence). Les résultats sont indépendants du choix effectué.

2 **Les aspects généraux concernant la qualité d'un modèle, l'interprétation des coefficients ... seront traités après la présentation de la régression logistique multiple**

3. Comparaison des modèles utilisant les fonctions de lien logit et probit.

- L'estimation du paramètre β_1 obtenue avec la fonction de lien logit est environ $\pi/\sqrt{3}$ fois plus grande que celle obtenue avec la fonction de lien probit. **Les estimations standardisées sont donc assez proches l'une de l'autre.**

- **Les résultats des tests de validité des modèles sont équivalents**.

- Enfin la comparaison des probabilités estimées montre que **les prévisions sont relativement similaires.**

- **PROBIT**

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	138.663	111.499	.
SC	141.268	116.709	.
-2 LOG L	136.663	107.499	29.164 with 1 DF (p=0.0001)
Score	.	.	26.399 with 1 DF (p=0.0001)

LOGIT

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	138.663	111.353	.
SC	141.268	116.563	.
-2 LOG L	136.663	107.353	29.310 with 1 DF (p=0.0001)
Score	.	.	26.399 with 1 DF (p=0.0001)

PROBIT

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate
INTERCPT	1	-3.1457	0.6246	25.3657	0.0001	.
AGE	1	0.0658	0.0133	24.3095	0.0001	0.771313

LOGIT

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	-5.3095	1.1337	21.9350	0.0001	.	.
AGE	1	0.1109	0.0241	21.2541	0.0001	0.716806	1.117

EXEMPLES DE PREVISIONS

- **PROBIT**

OBS	ID	AGRP	AGE	CHD	_LEVEL_	PCALCULE	INF95	SUP95
1	1	1	20	0	1	0.03365	0.00534	0.13436
39	39	4	40	1	1	0.30378	0.20511	0.41932
91	91	8	60	0	1	0.78888	0.63227	0.89745
100	100	8	69	1	1	0.91846	0.76581	0.98051

- **LOGIT**

OBS	ID	AGRP	AGE	CHD	_LEVEL_	PCALCULE	INF95	SUP95
1	1	1	20	0	1	0.04348	0.01207	0.14470
39	39	4	40	1	1	0.29471	0.19510	0.41873
91	91	8	60	0	1	0.79344	0.63245	0.89556
100	100	8	69	1	1	0.91246	0.76287	0.97124