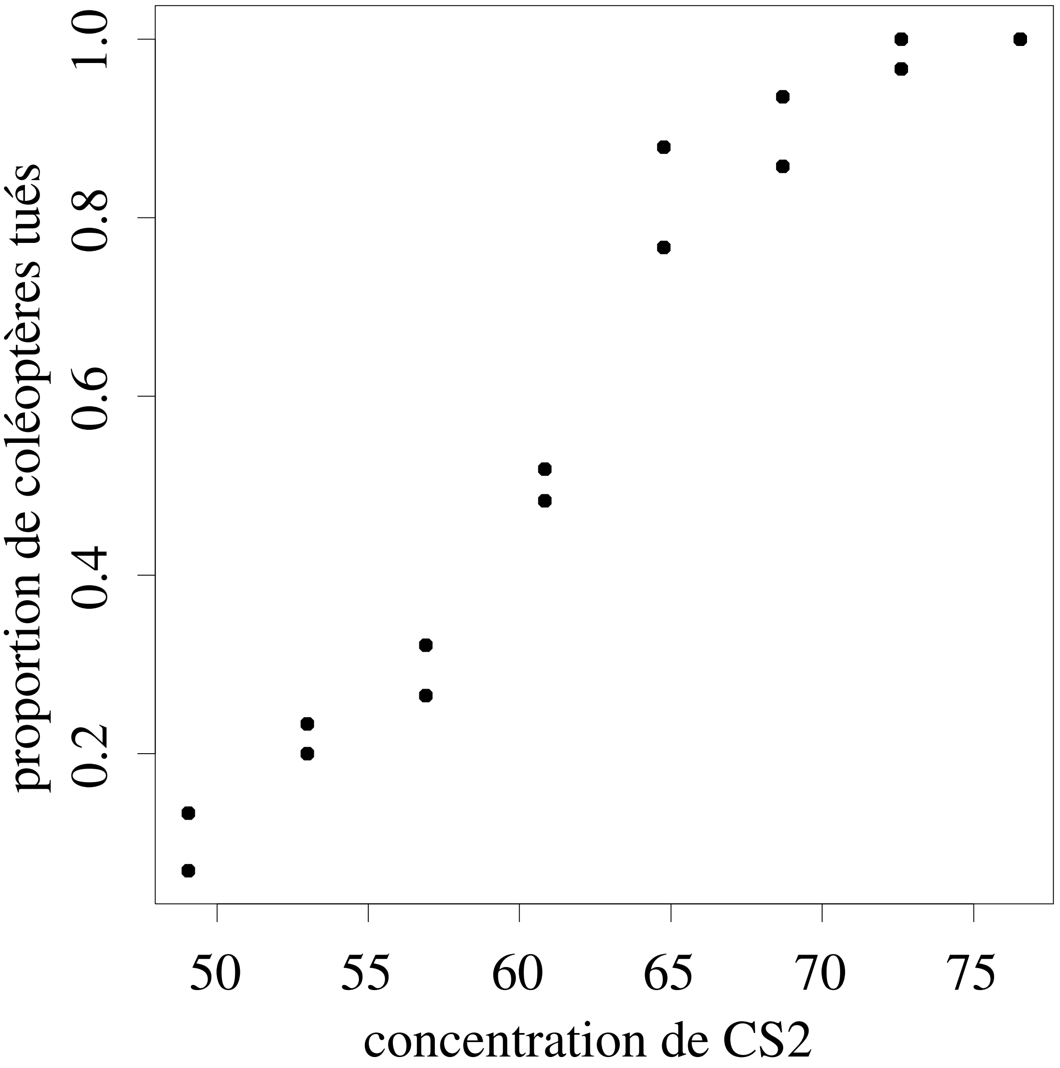
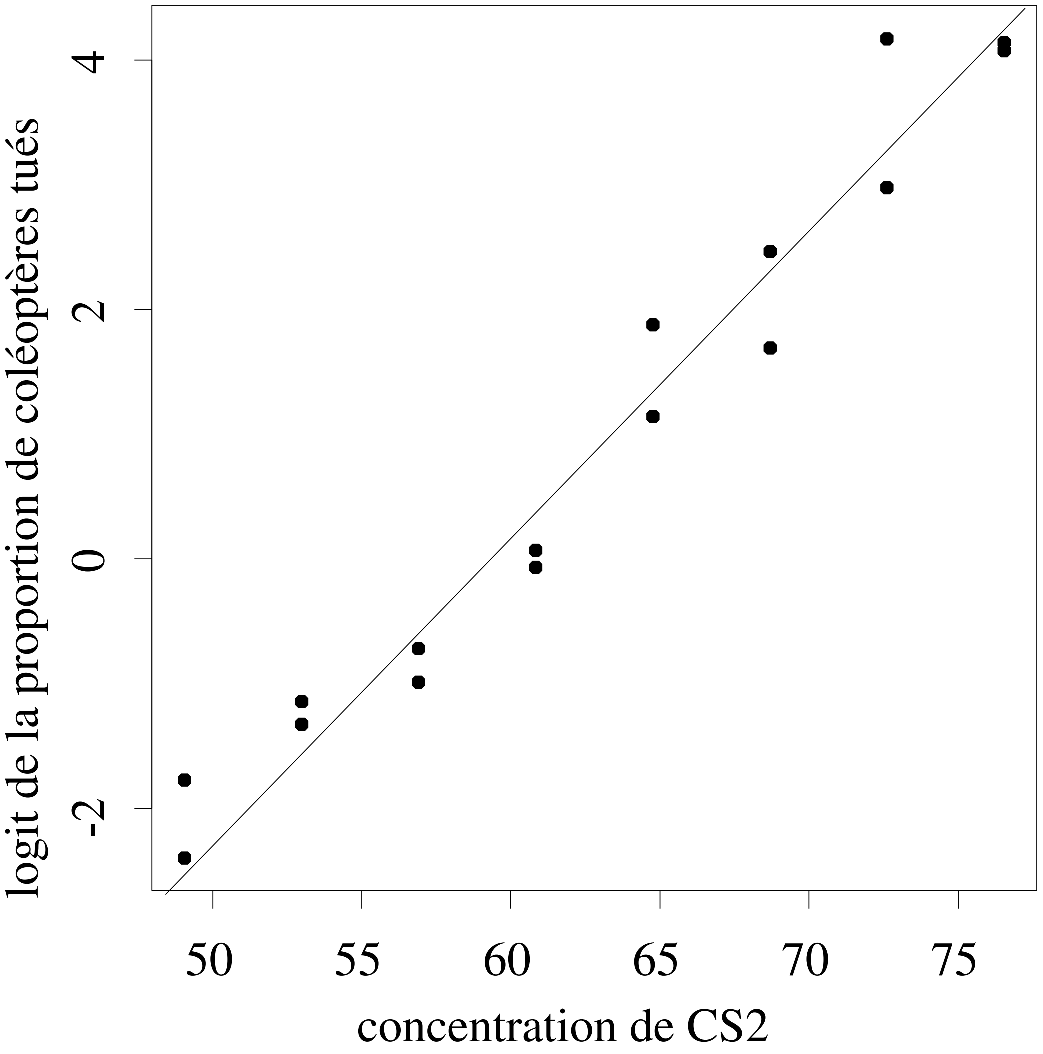
MORTALITÉ D'UN COLÉOPTÈRE DE LA FARINE AU CS2

concentration de CS2 (mg/l)	répétition 1			répétition 2		
cc	y	n	$p = y/n$	y	n	$p = y/n$
49.06	2	29	0.069	4	30	0.133
52.99	7	30	0.233	6	30	0.200
56.91	9	28	0.321	9	34	0.265
60.84	14	27	0.519	14	29	0.483
64.76	23	30	0.767	29	33	0.879
68.69	29	31	0.935	24	28	0.857
72.61	29	30	0.967	32	32	1.000
76.54	29	29	1.000	31	31	1.000

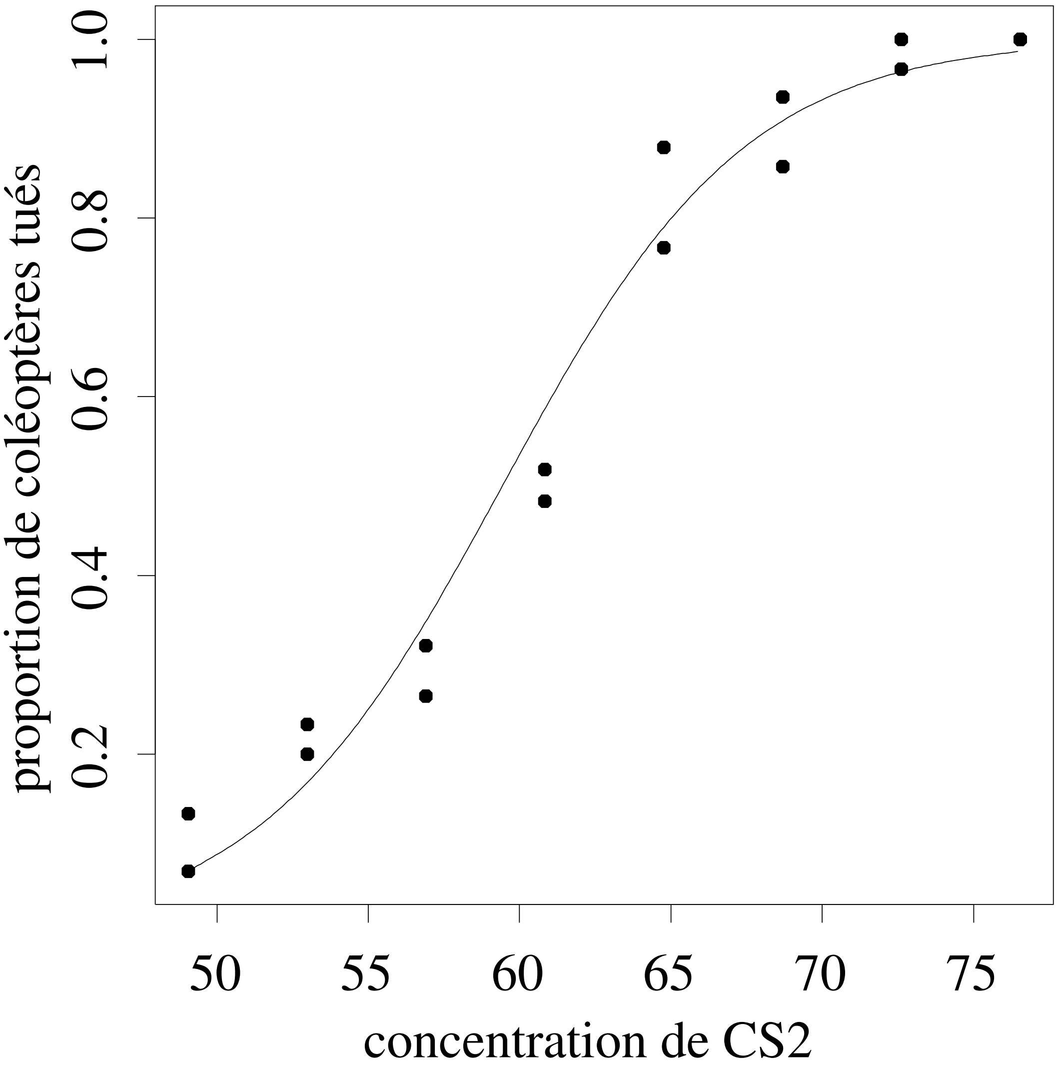
ÉCHELLE D'ORIGINE



ÉCHELLE LOGIT



RETOUR À L'ÉCHELLE D'ORIGINE



LES RÉSIDUS

Résidus *bruts* $y_i - \widehat{y}_i$

Résidus de *Pearson*

$$R_{P_i} = \frac{y_i - \widehat{y}_i}{\text{écart-type}(y_i)} = \frac{y_i - n_i\widehat{p}_i}{\sqrt{n_i\widehat{p}_i(1-\widehat{p}_i)}}$$

Résidus de *Pearson standardisés*

$$r_{P_i} = \frac{y_i - \widehat{y}_i}{\text{écart-type}(y_i - \widehat{y}_i)} = \frac{y_i - n_i\widehat{p}_i}{\sqrt{n_i\widehat{p}_i(1-\widehat{p}_i)(1-h_i)}} = \frac{R_{P_i}}{\sqrt{1-h_i}}$$

Résidus de *Déviance*

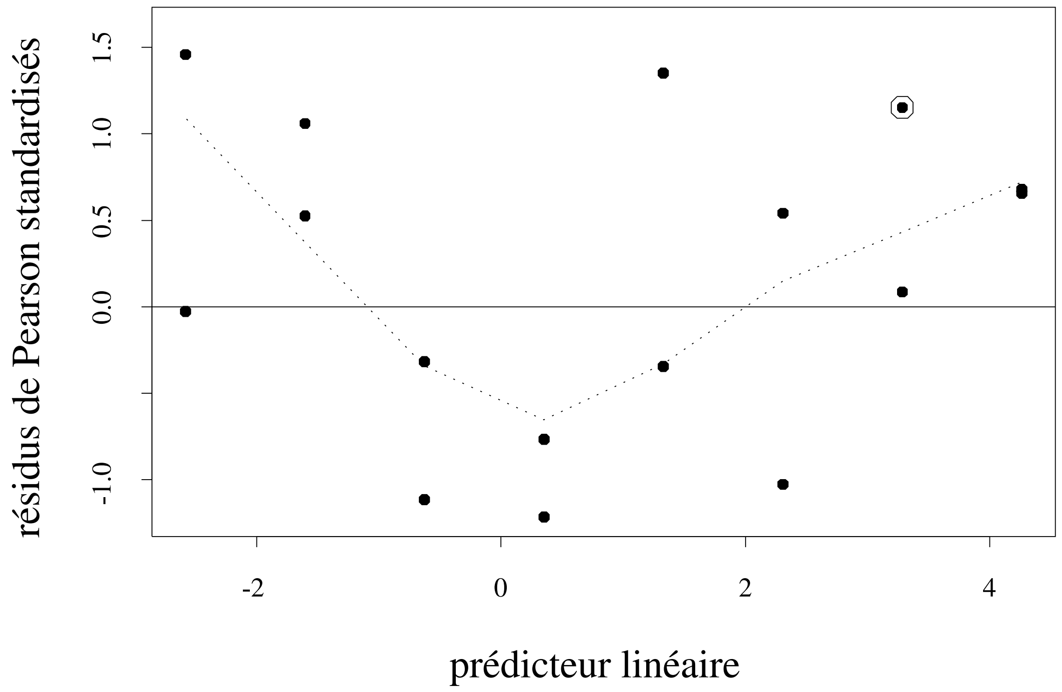
$$R_{D_i} = signe(y_i - \widehat{y}_i)\sqrt{2y_i \log\frac{y_i}{\widehat{y}_i} + 2(n_i - y_i)\log\frac{n_i - y_i}{n_i - \widehat{y}_i}}$$

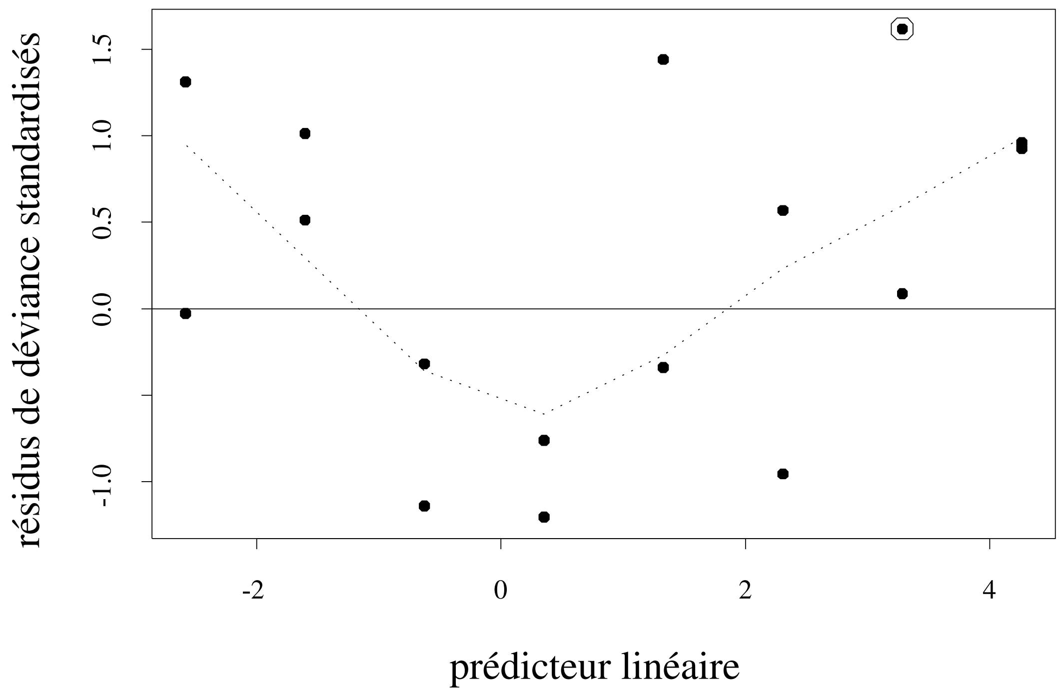
Résidus de *Déviance standardisés*

$$r_{D_i} = \frac{R_{D_i}}{\sqrt{1-h_i}}$$

$$\chi^2_{\text{Pearson}} = \sum_i R^2_{P_i} \qquad D = \sum R^2_{D_i}$$

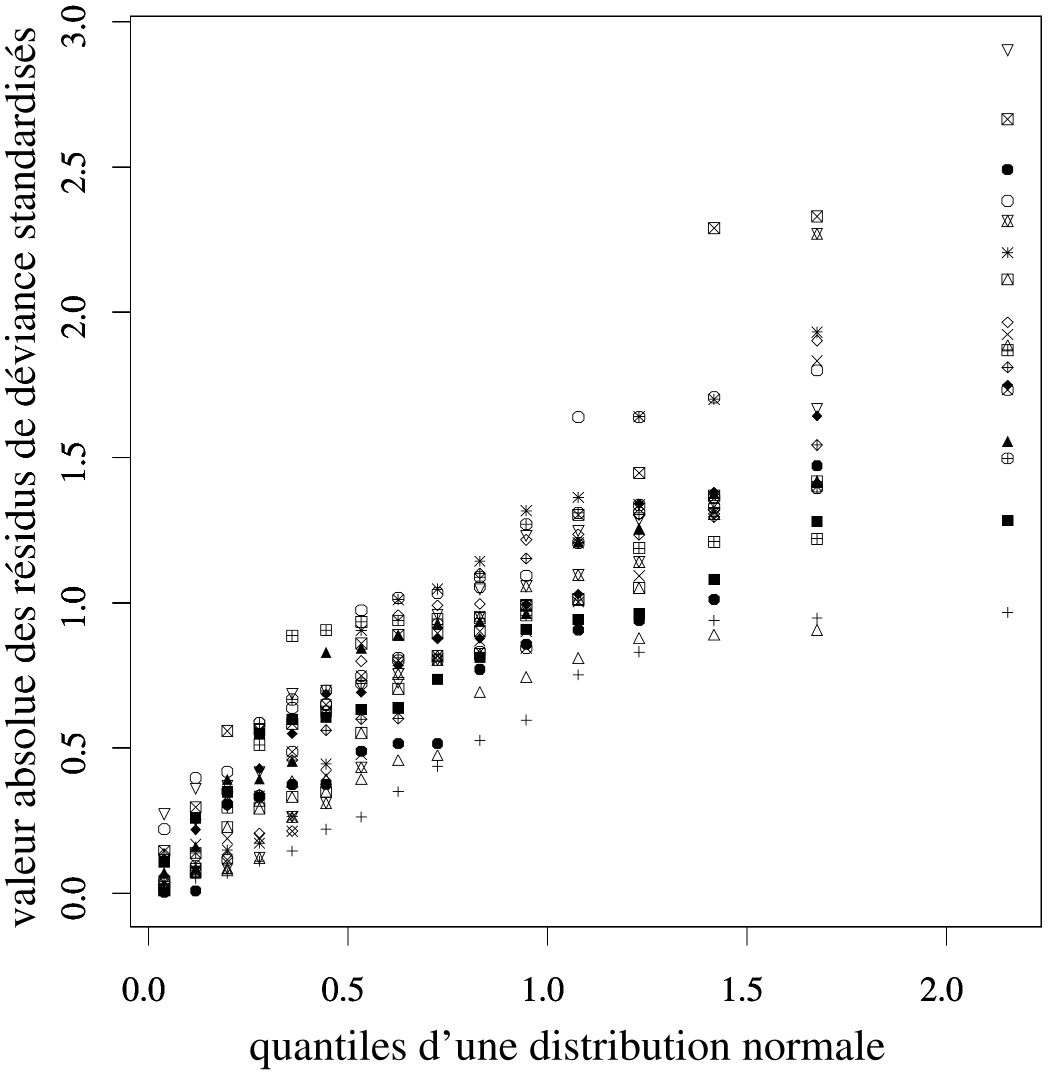
REPRÉSENTATION DES RÉSIDUS





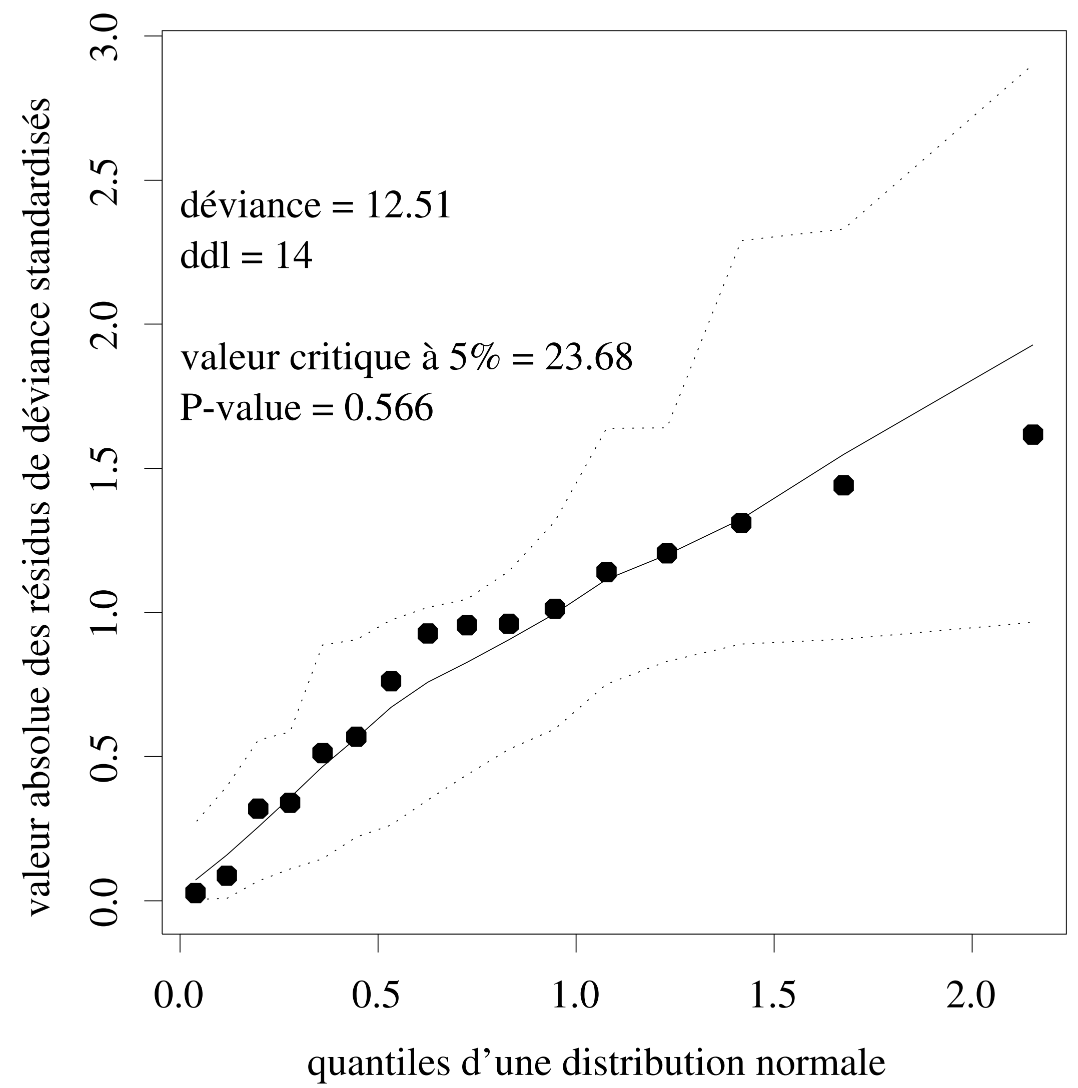
ADÉQUATION DU PRÉDICTEUR LINÉAIRE

half-normal plot pour
19 échantillons simulés



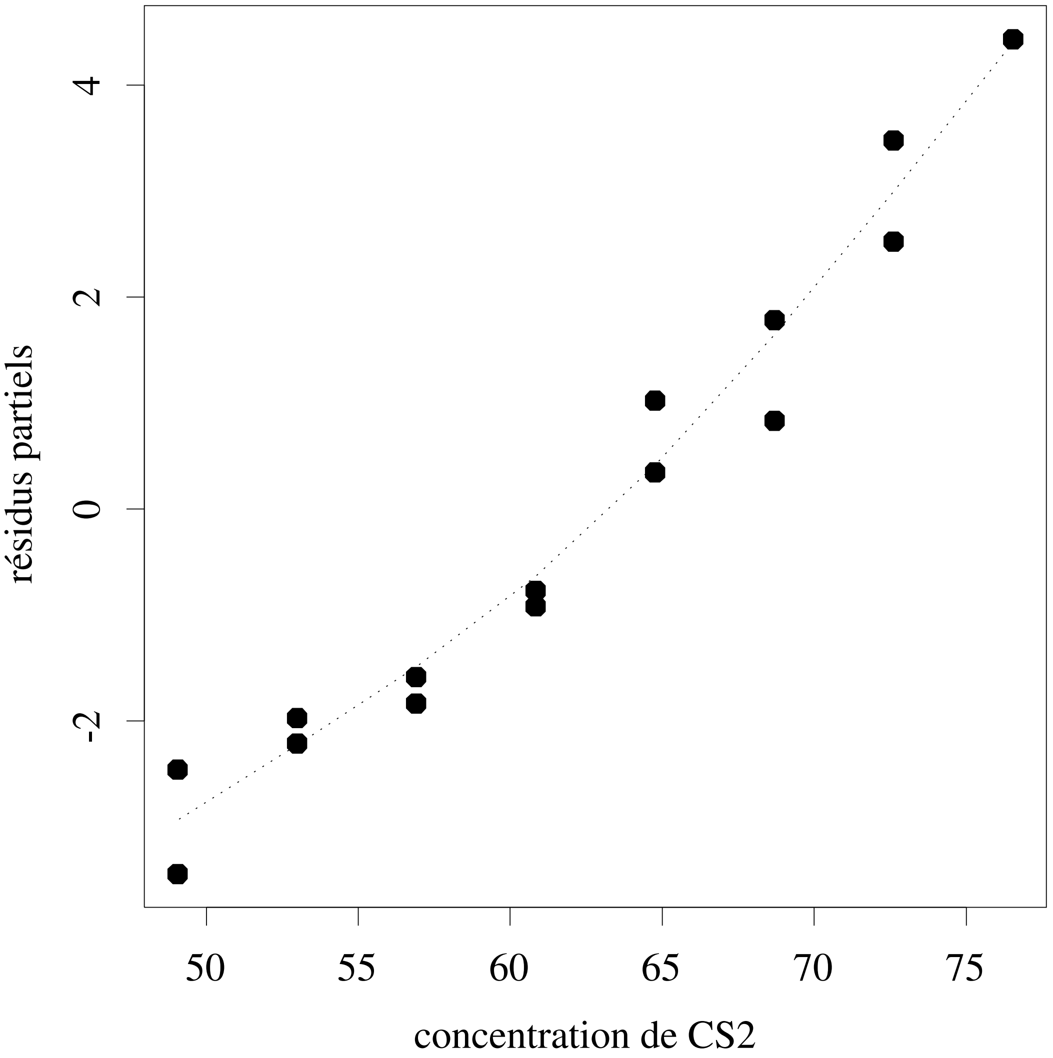
ADÉQUATION DU PRÉDICTEUR LINÉAIRE

half-normal plot et enveloppe simulée

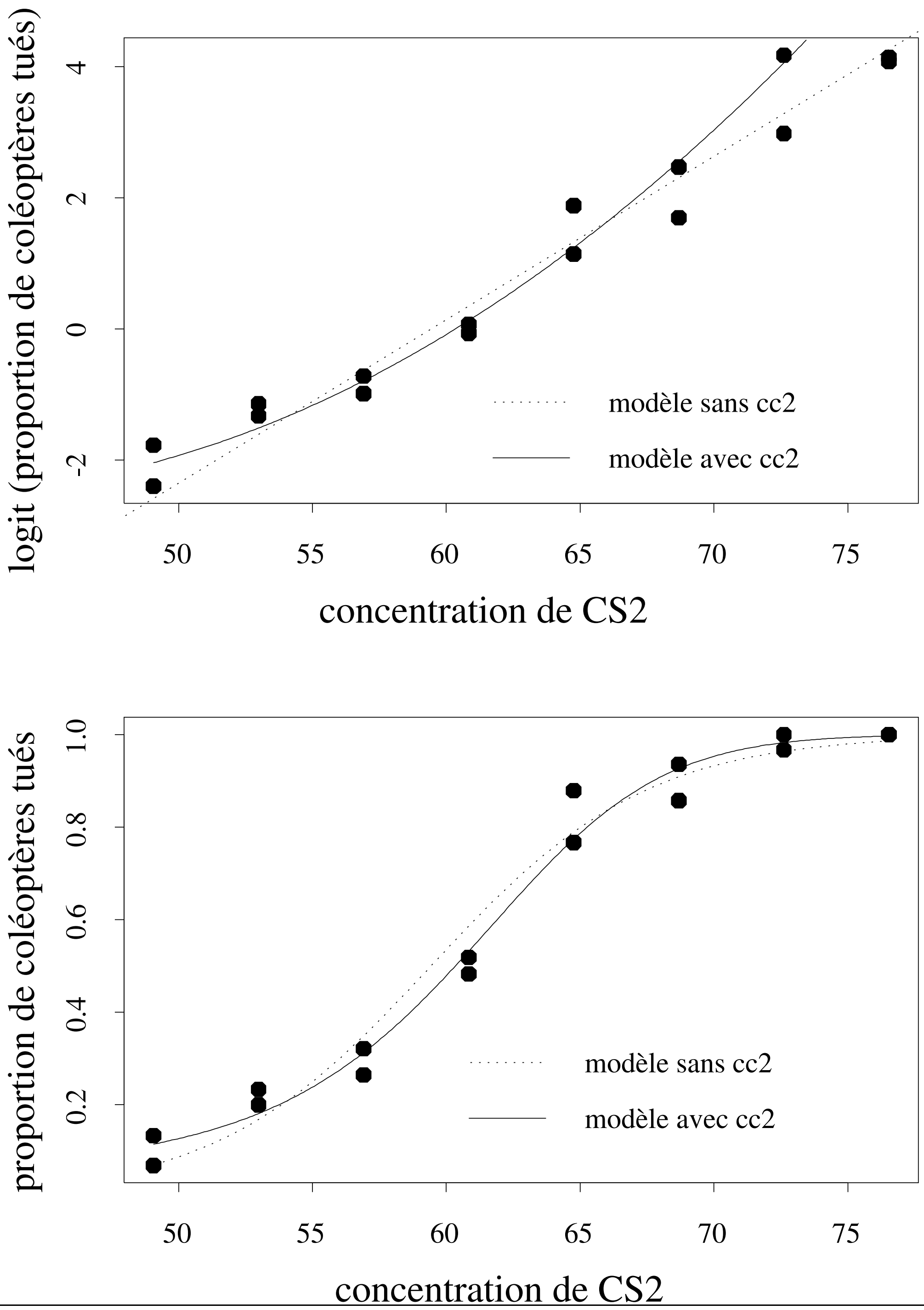


ADÉQUATION DU PRÉDICTEUR LINÉAIRE

graphique des résidus partiels pour cc

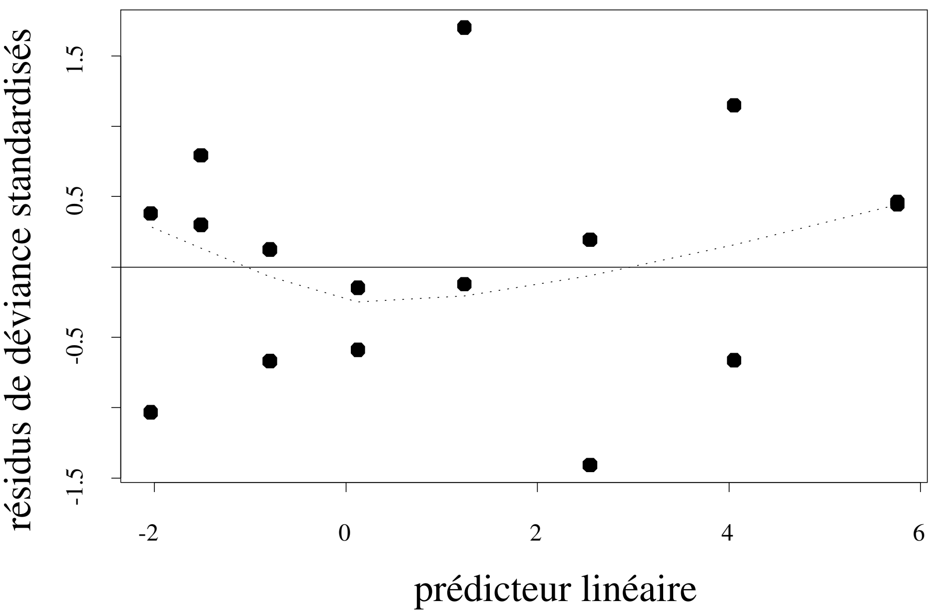


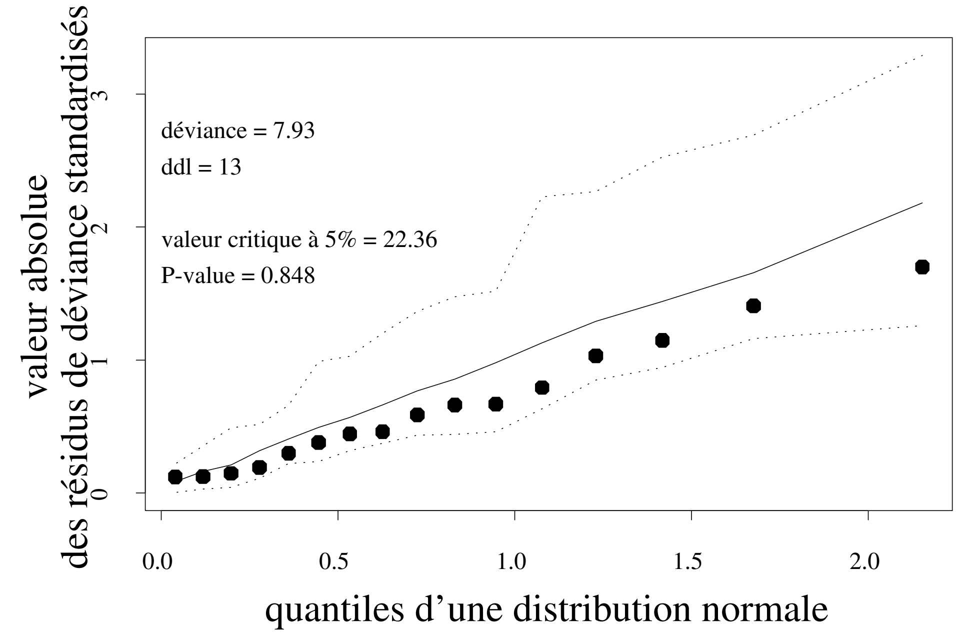
INTRODUCTION D'UN TERME QUADRATIQUE



ADÉQUATION DU MODÈLE

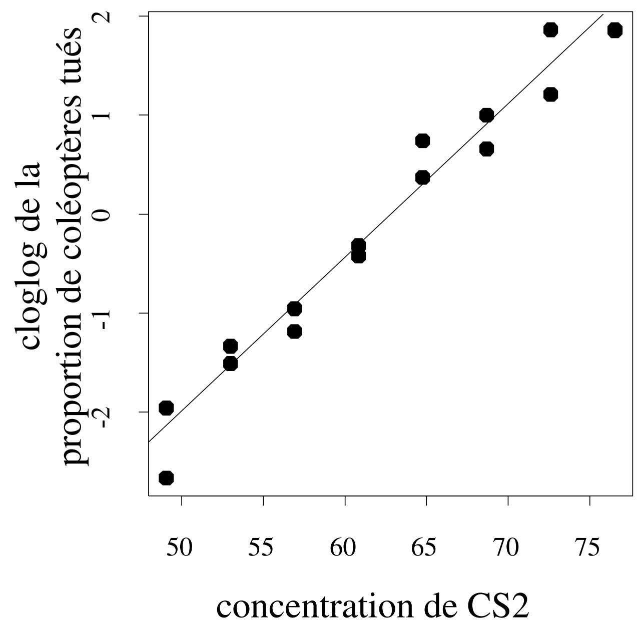
Modèle quadratique

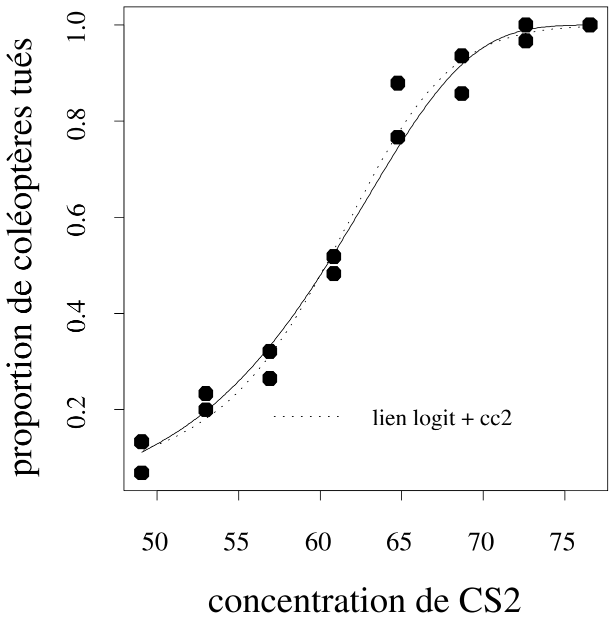




CHOIX DU LIEN CLOGLOG

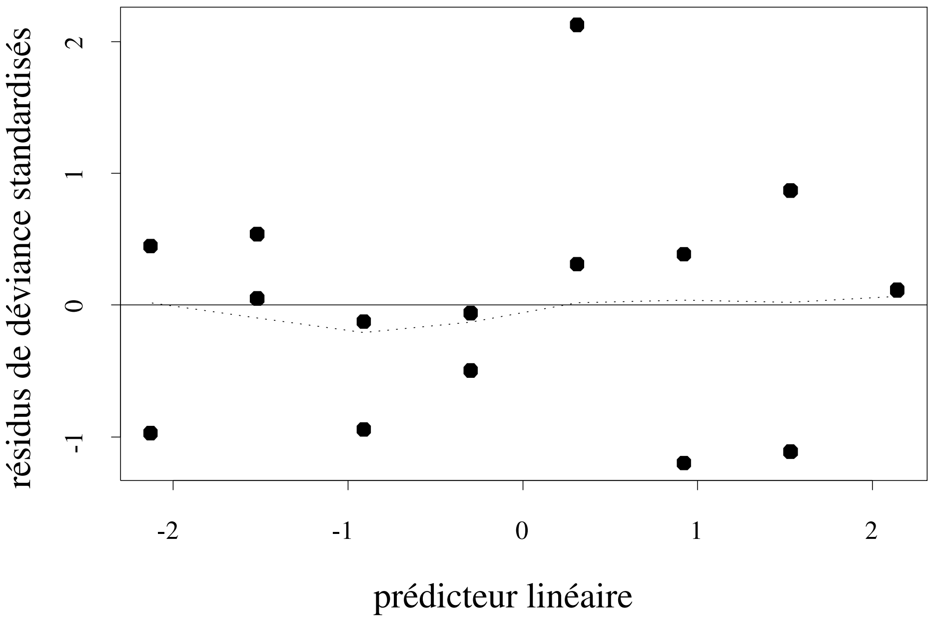
Modèle de degré 1

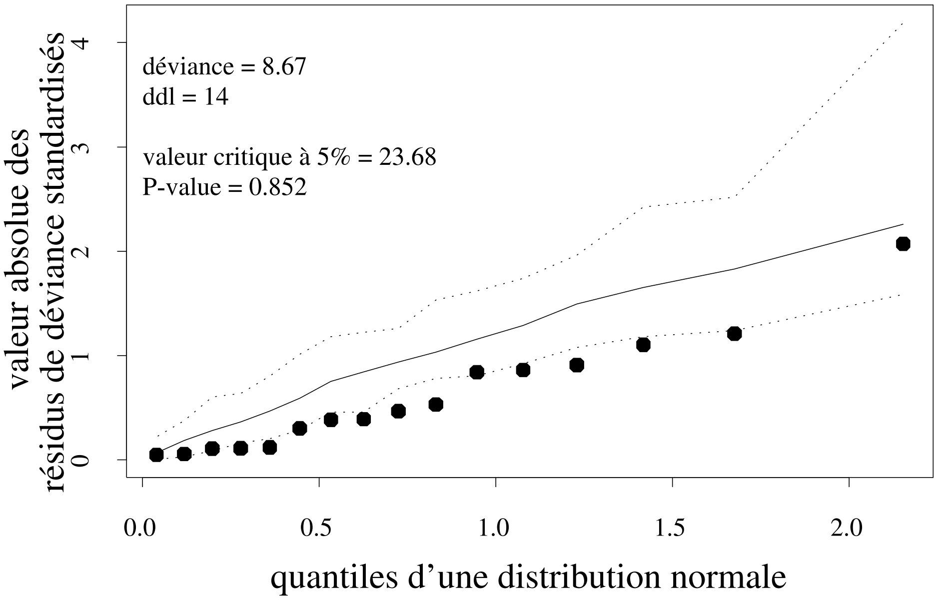




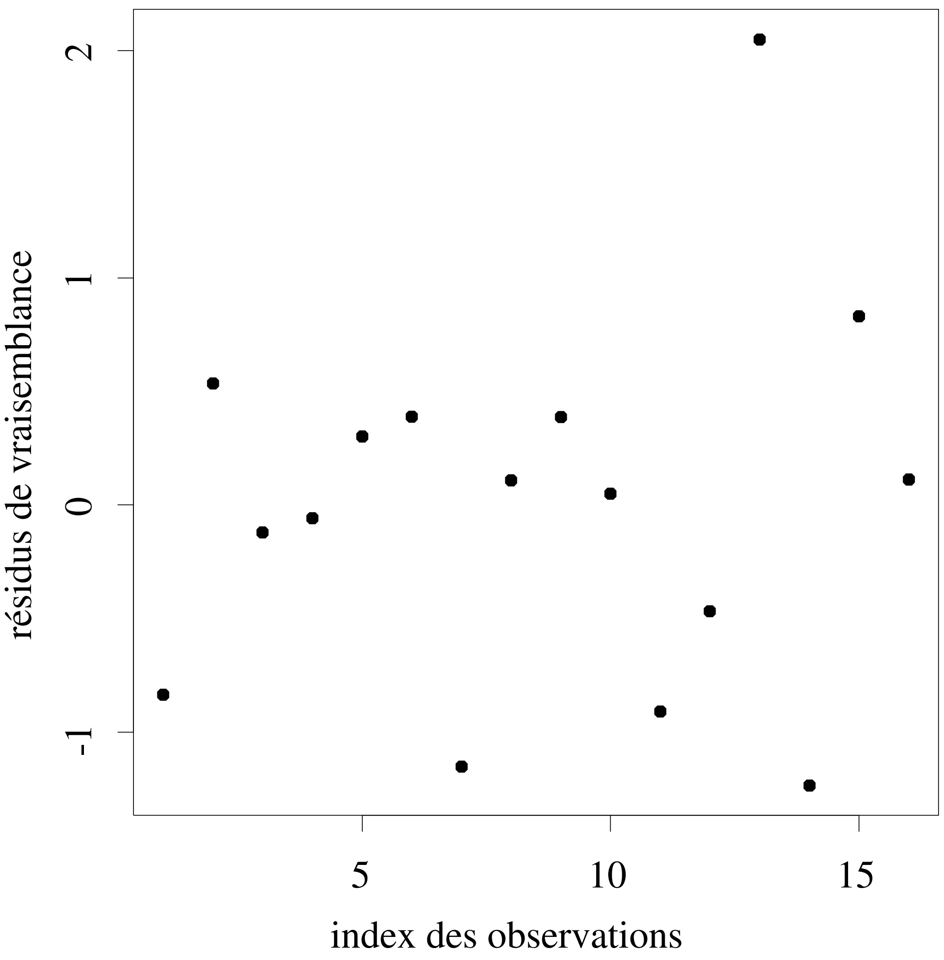
VALIDATION DU MODÈLE

Modèle de degré 1, lien cloglog





IDENTIFICATION DES POINTS EXTÉRIEURS

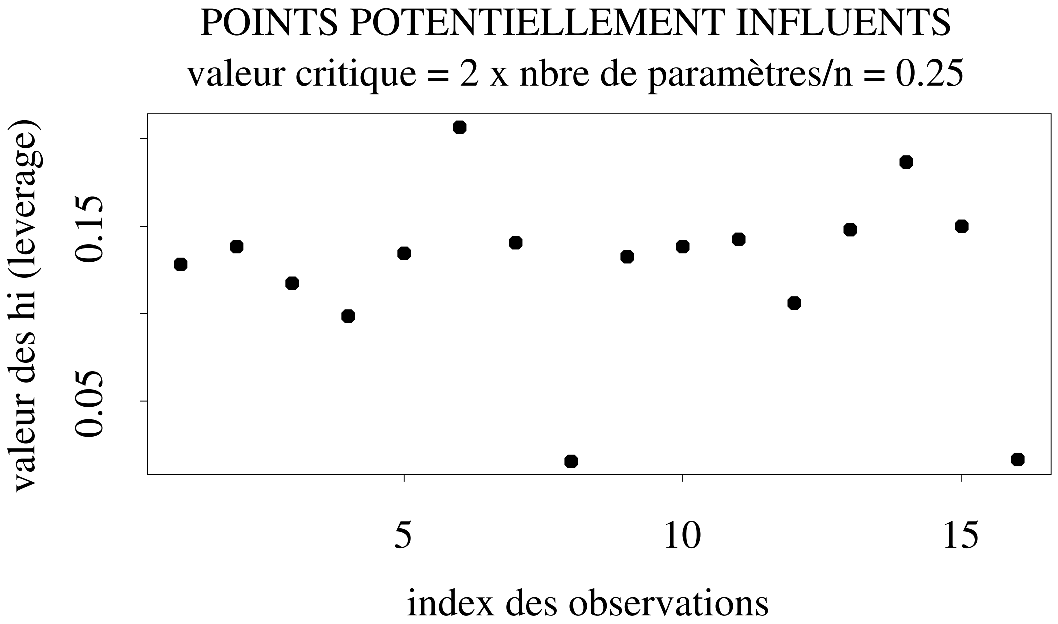


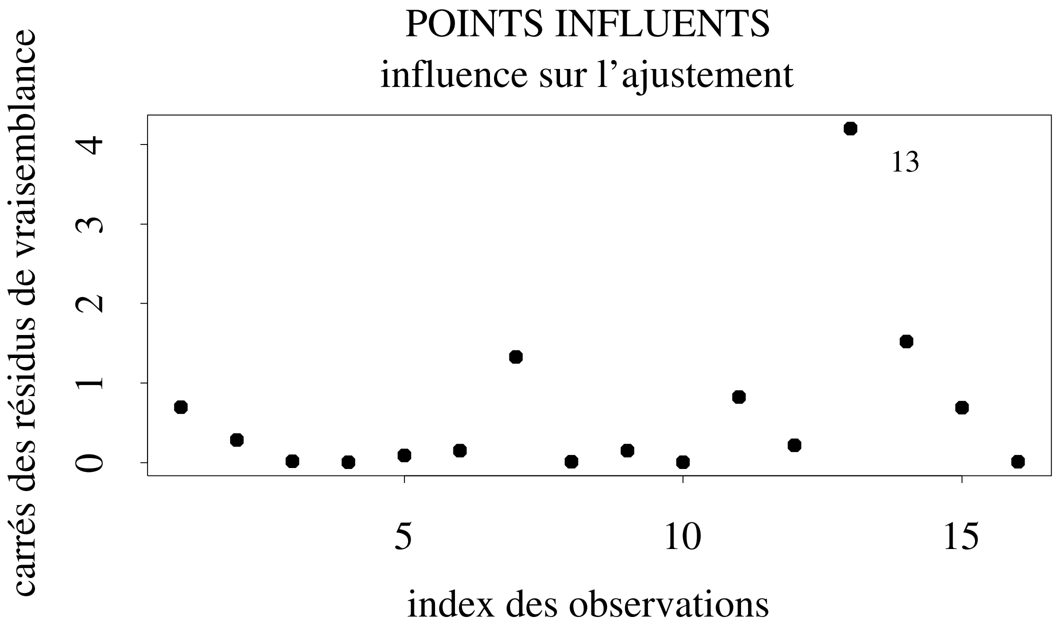
Test de discordance

i e obs. discordante si $r_{L_i} > \phi^{-1}(1 - \frac{\alpha}{2n})$

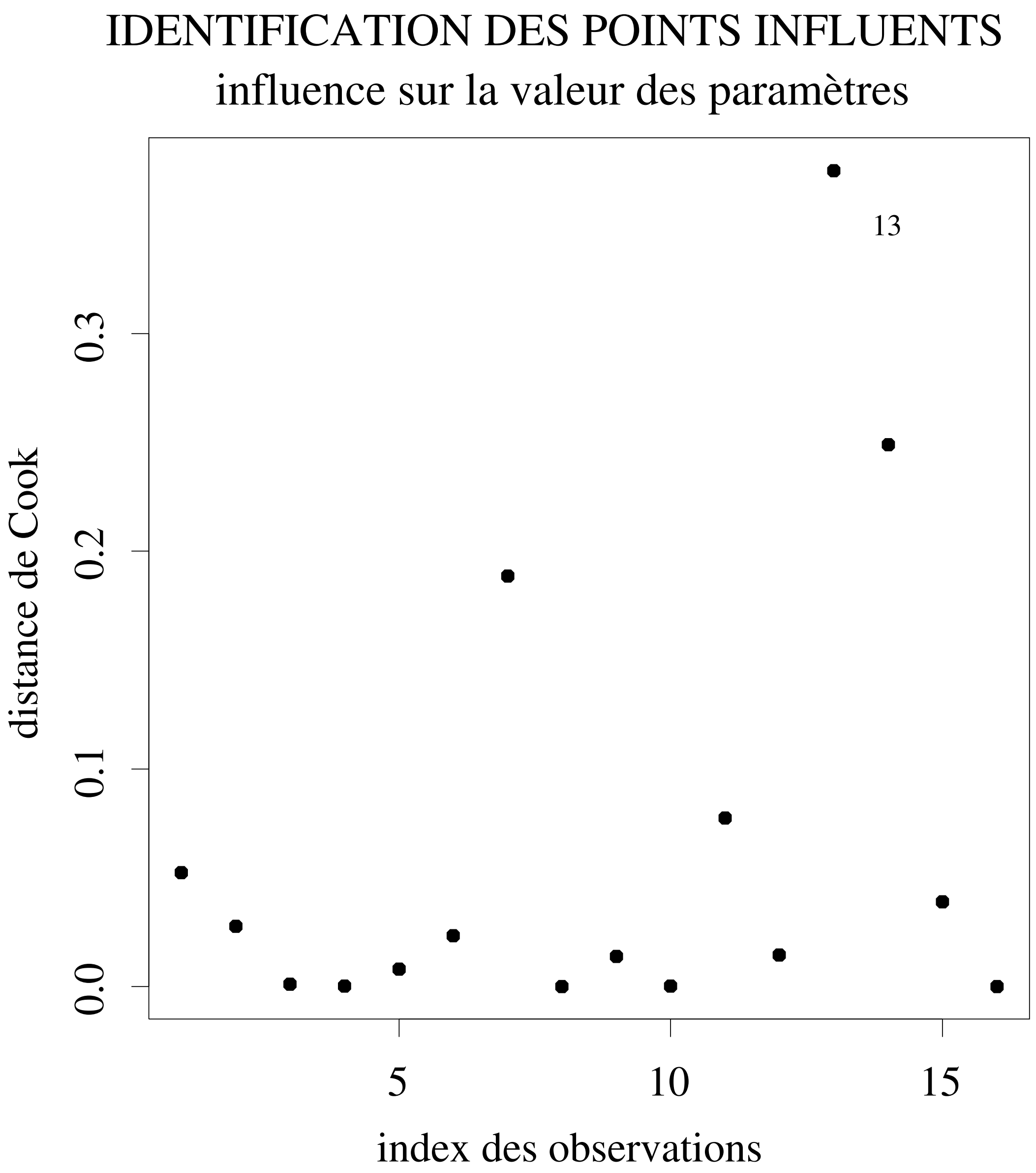
valeur critique au seuil 5% : 2.96

IDENTIFICATION DES POINTS INFLUENTS

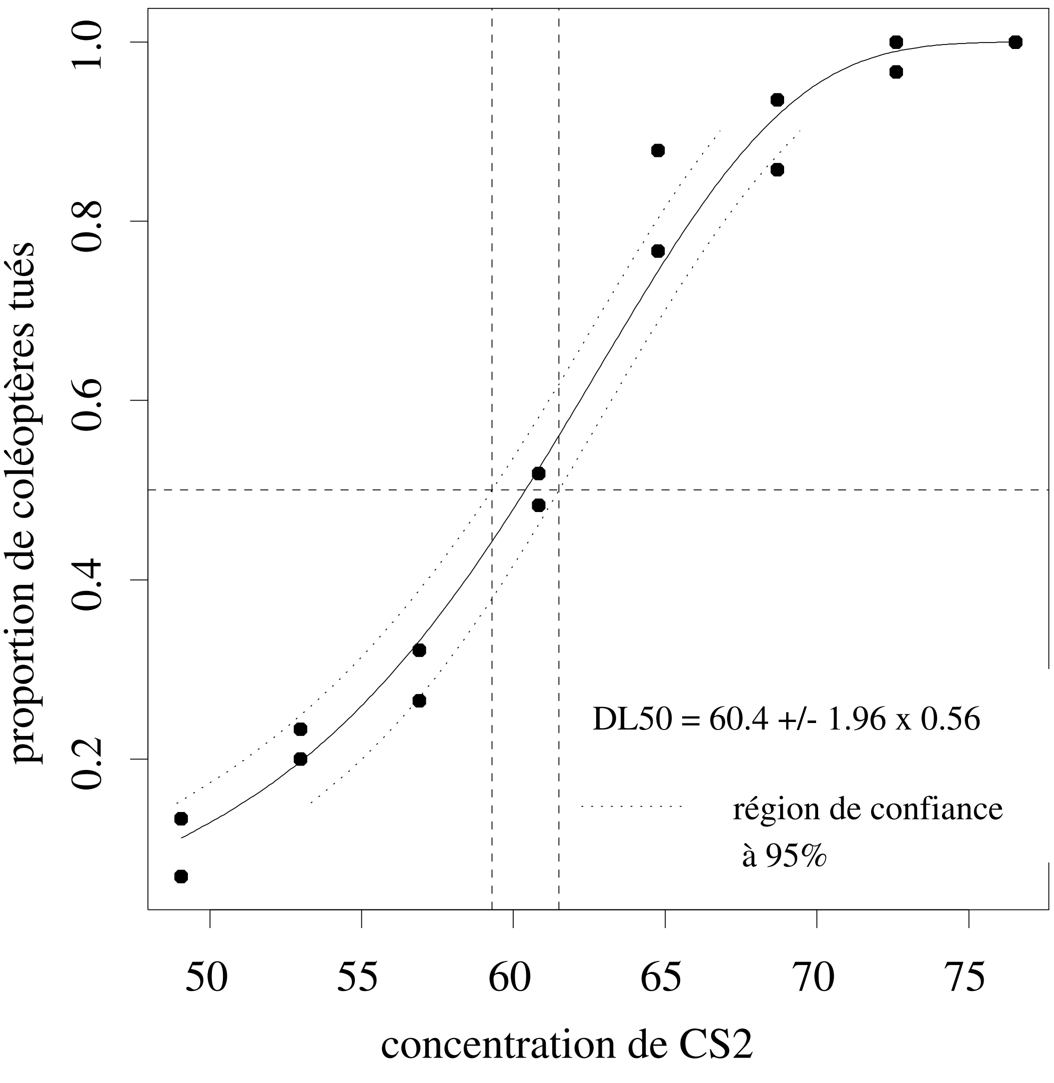




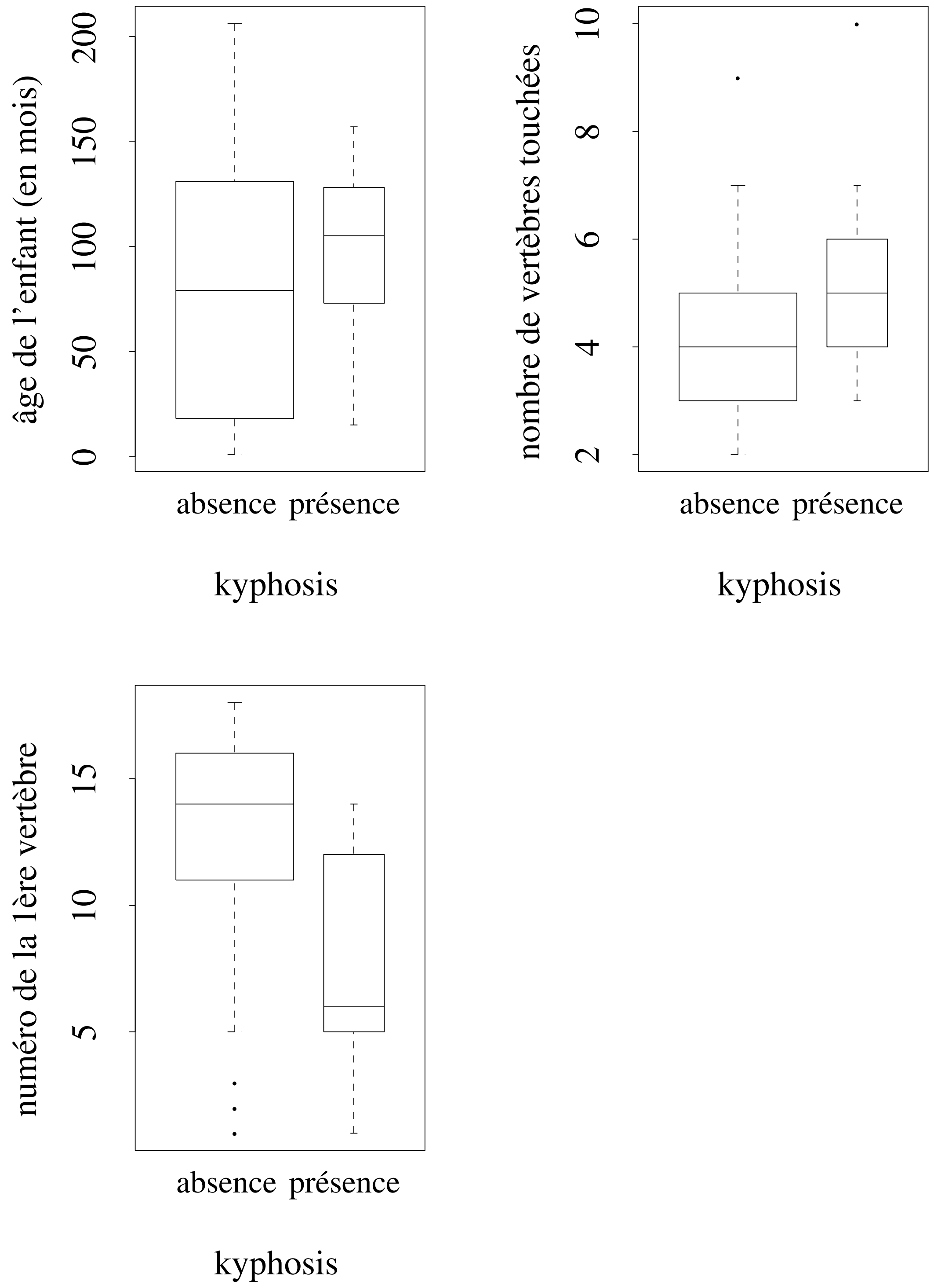
IDENTIFICATION DES POINTS INFLUENTS



Modèle ajusté, DL_{50}



DONNÉES KYPHOSIS



MODÈLE ADDITIF, LIEN LOGIT

ESTIMATION DES PARAMÈTRES :

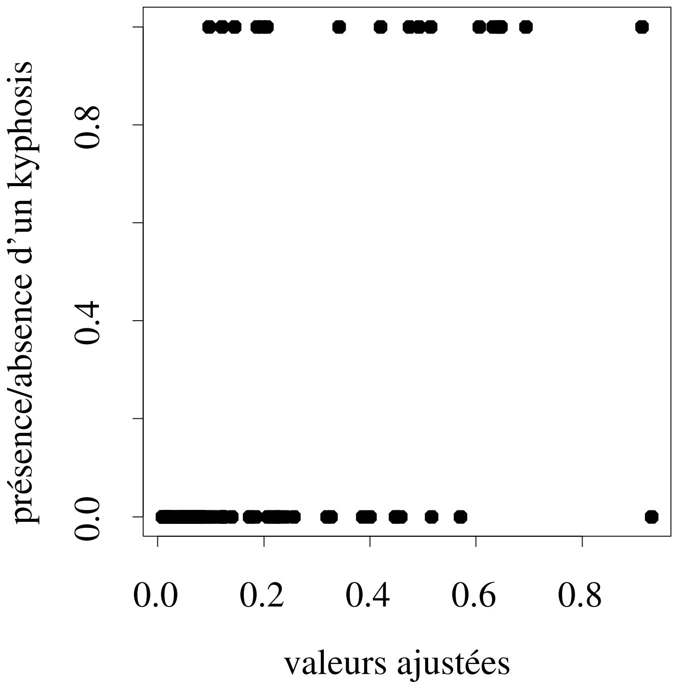
	valeur	écart-type	valeur du T
constante	-2.037	1.449	-1.406
numéro	-0.207	0.068	-3.051
nombre	0.411	0.225	1.827
âge	0.011	0.006	1.696

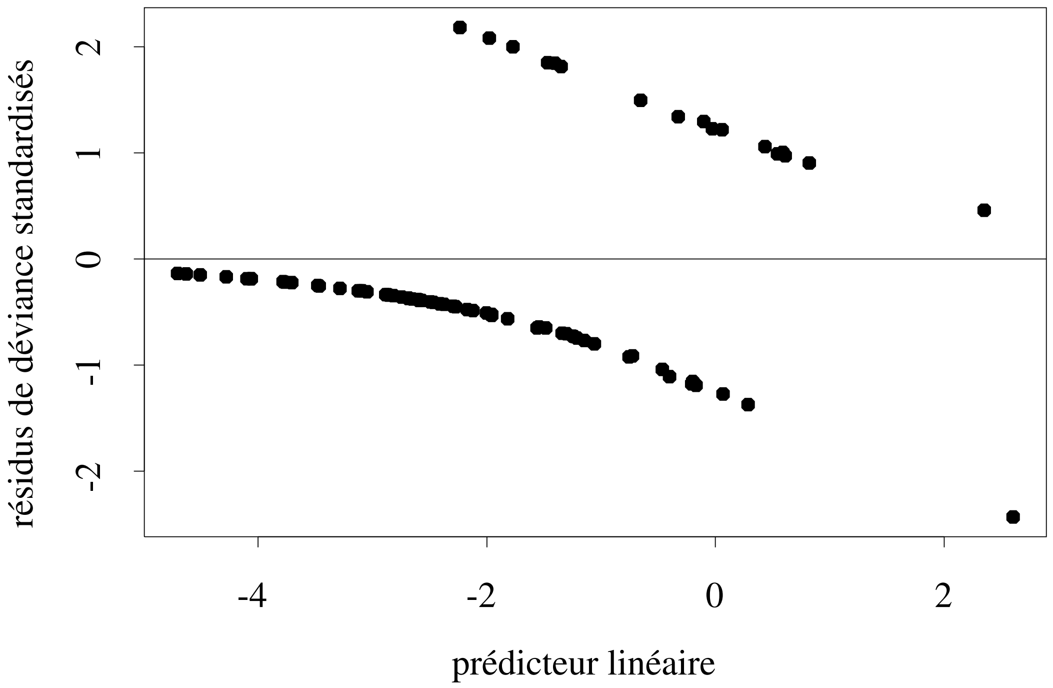
DÉVIANCE DU MODELE NUL : 83.23 pour 80 ddl
DÉVIANCE RÉSIDUELLE : 61.38 pour 77 ddl

TABLE DE DÉVIANCE :

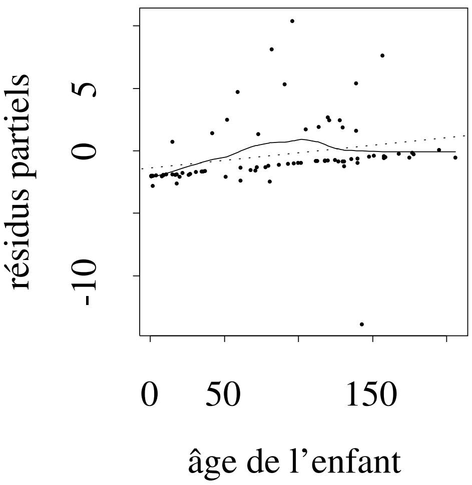
	ddl	déviance	dl résiduels	déviance résiduelle	Pr(Chi)
modèle nul			80	83.23	
numéro	1	15.16	79	68.07	0.0001
nombre	1	3.53	78	64.54	0.0601
âge	1	3.16	77	61.38	0.0756

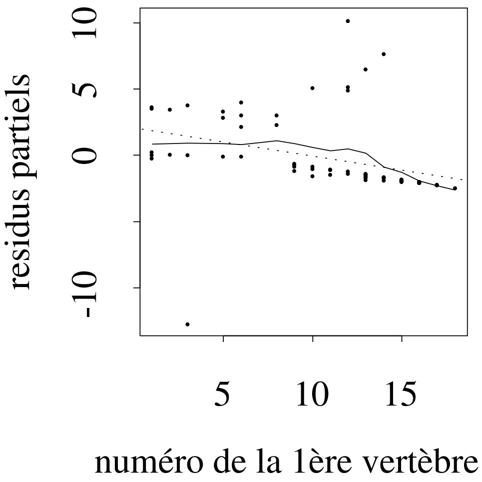
VALIDATION DE MODÈLE

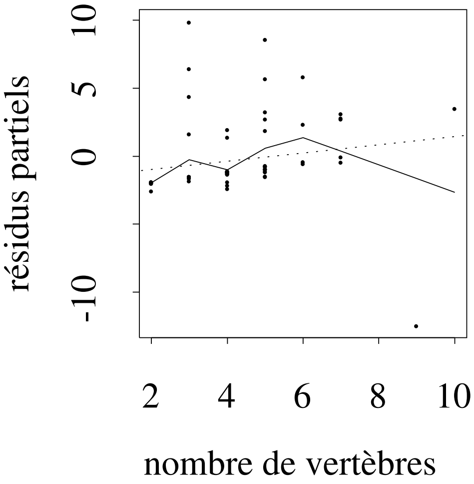




VALIDATION DU MODÈLE







MODÈLE QUADRATIQUE

lien logit

TABLE DE DÉVIANCE :

	ddl	déviance	ddl résiduels	déviance résiduelle	Pr(Chi)
modèle nul			80	83.23	
poly(numéro,2)	2	21.53	78	61.70	0.00002
poly(nombre,2)	2	2.19	76	59.51	0.33363
poly(âge,2)	2	11.67	74	47.84	0.00292

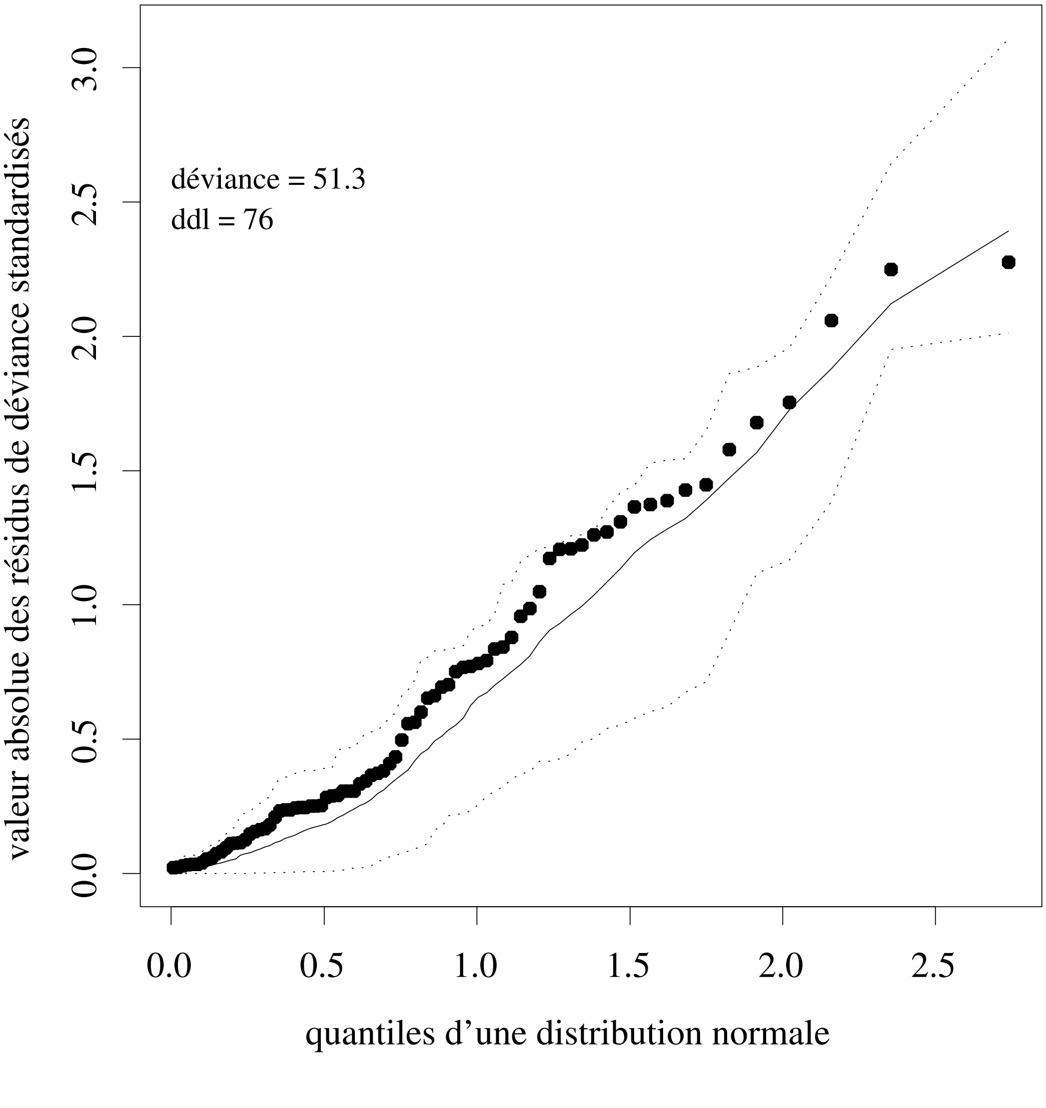
SOUS-MODÈLE

TABLE DE DÉVIANCE :

	ddl	déviance	ddl résiduels	déviance résiduelle	Pr(Chi)
modèle nul			80	83.23	
poly(numéro,2)	2	21.53	78	61.70	0.00002
poly(âge,2)	2	10.40	76	51.30	0.00551

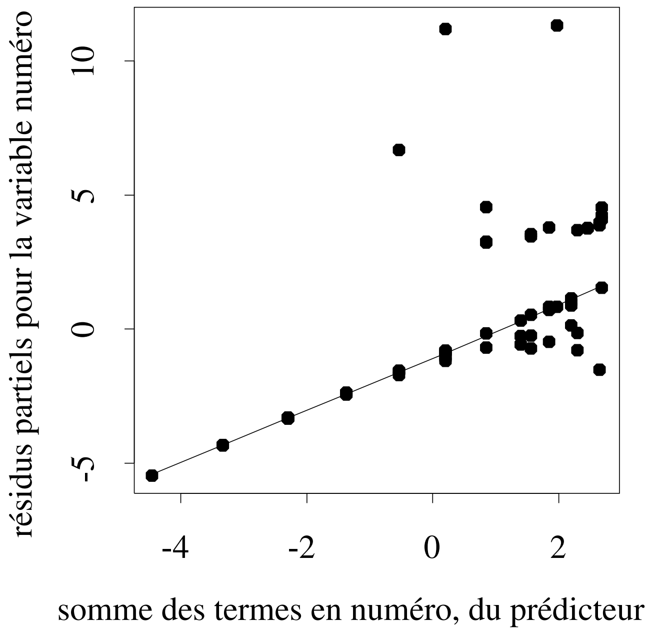
VALIDATION DE MODÈLE

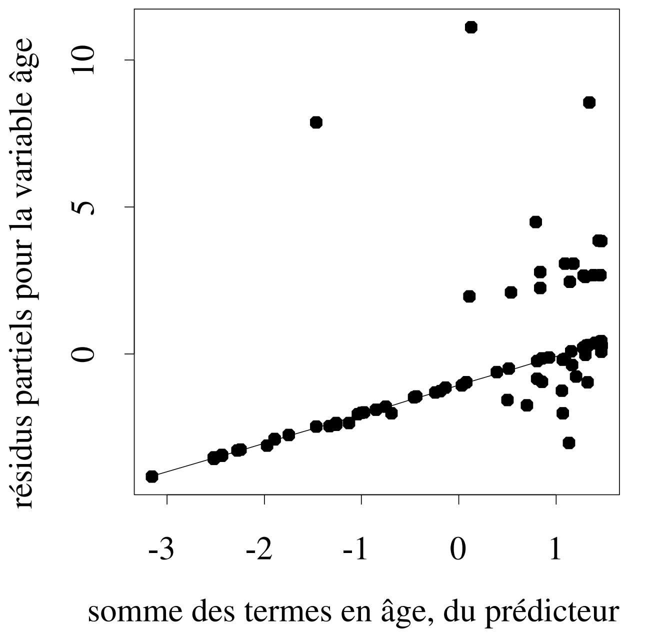
Modèle quadratique en muméro et en âge



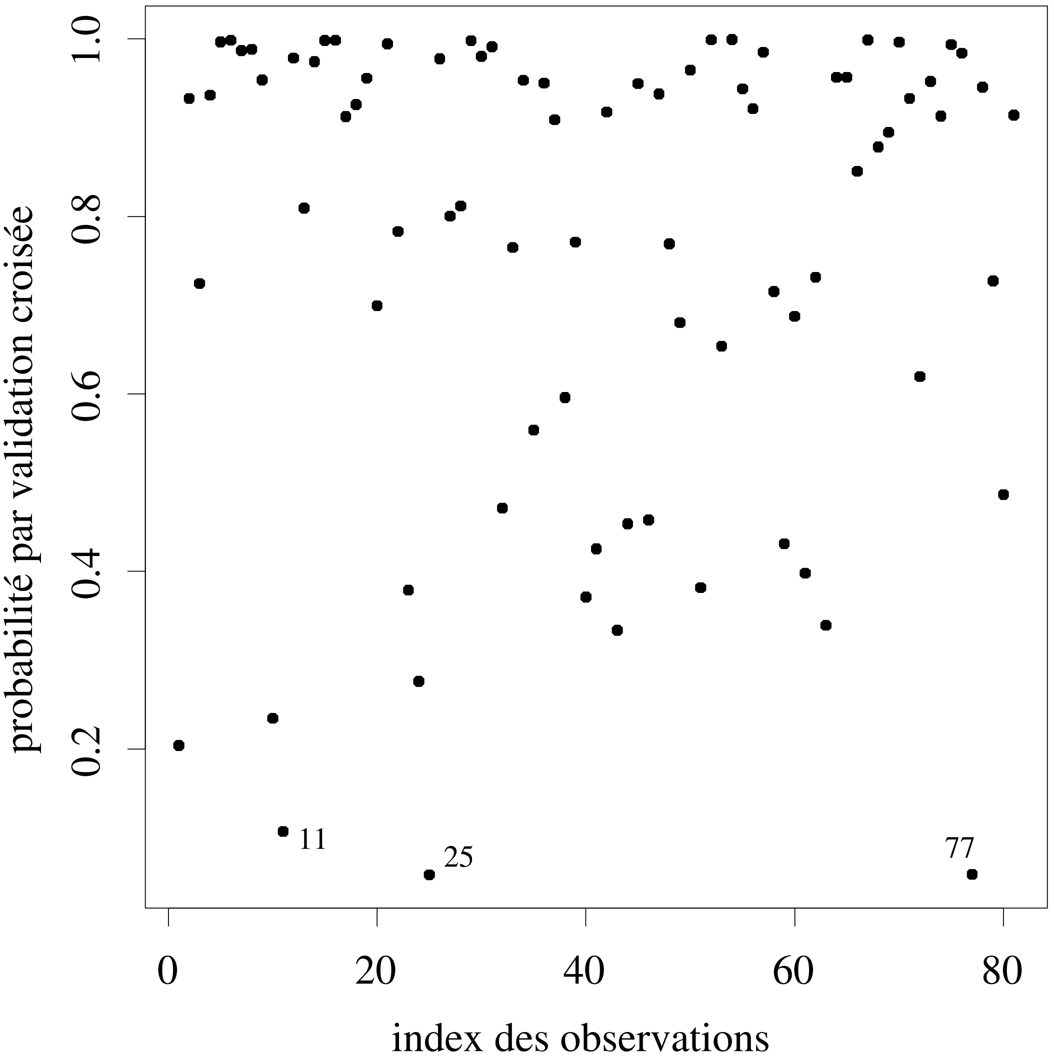
VALIDATION DE MODÈLE

Modèle quadratique en muméro et en âge



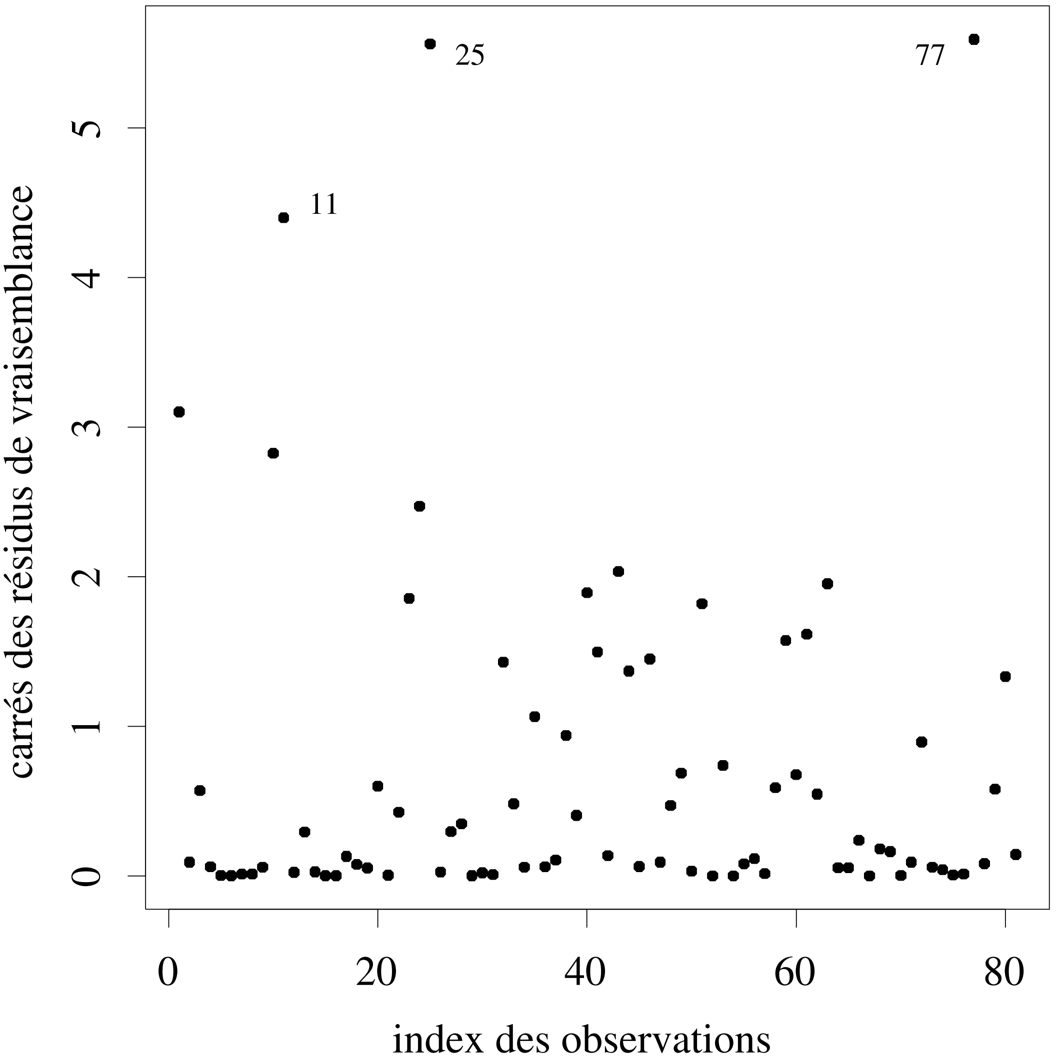


IDENTIFICATION DES POINTS EXTÉRIEURS



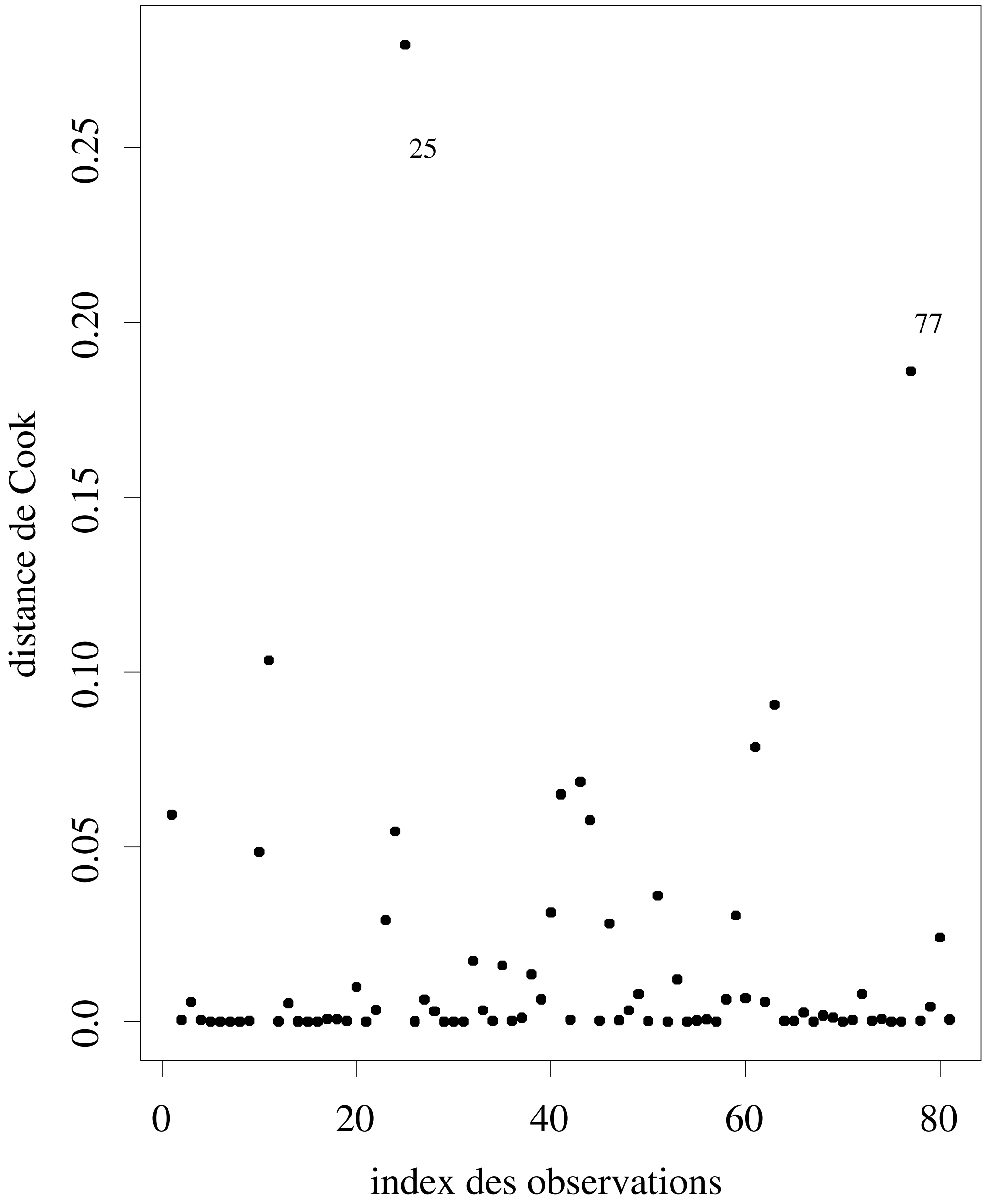
IDENTIFICATION DES POINTS INFLUENTS

influence sur l'ajustement



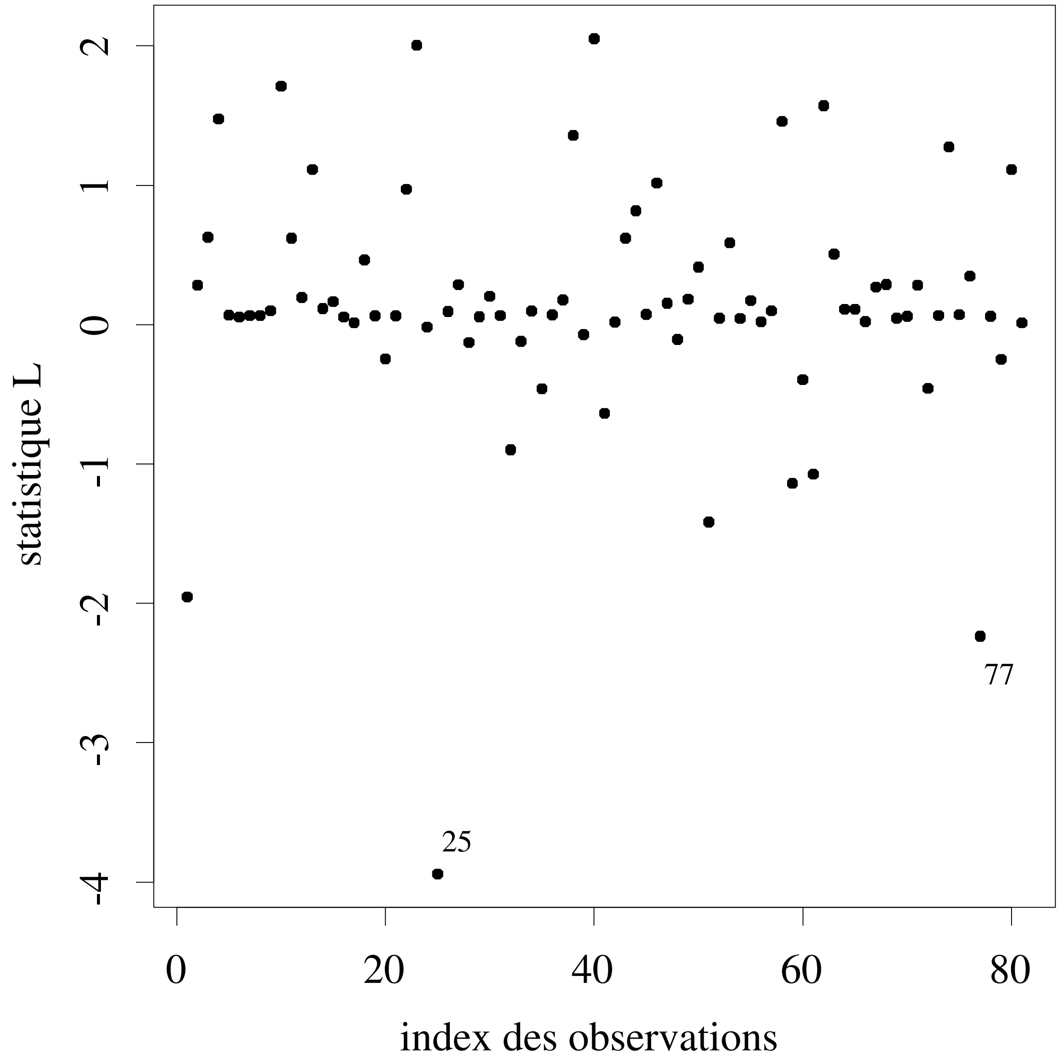
IDENTIFICATION DES POINTS INFLUENTS

influence sur les paramètres



IDENTIFICATION DES POINTS INFLUENTS

influence sur la comparaison de 2 déviances



ADÉQUATION DU MODÈLE

outil de validation	**type de données**	**usage**
déviance résiduelle	groupées	permet de voir si le modèle est significativement différent du modèle saturé.
graphique des résidus de déviance standardisés en fonction du prédicteur linéaire	groupées	possibilité de meilleur modèle s'il apparaît une structure dans le graphique, mise en évidence de points extérieurs.
"half-normal plot" des résidus de déviance standardisés et enveloppe simulée	groupées non-groupées	possibilité de meilleur modèle si déviation systématique par rapport à la ligne moyenne, mise en évidence de points extérieurs.
graphique des résidus partiels en fonction des variables explicatives	groupées non-groupées	une tendance non linéaire dans le graphique suggère une transformation de la variable explicative et le type de transformation (quadratique, cubique...).
changement de fonction de lien	groupées non groupées	si on n'a pas d'idée sur la distribution des seuils de tolérance, on peut tester différentes fonctions de lien, et retenir le modèle le plus adéquat au regard de la parcimonie, de la déviance, des résidus...

ÉTUDE D'OBSERVATIONS PARTICULIÈRES

outil de validation	**type de données**	**usage**
graphique indexé des probabilités estimées par validation croisée	non groupées	identification de points extérieurs.
graphique indexé des "leverage"	groupées non groupées	mise en évidence de points potentiellement influents.
graphique indexé des carrés des résidus de vraisemblance test de discordance	groupées non groupées groupées	identification des observations ayant une influence sur la qualité de l'ajustement, identification de points extérieurs.
graphique indexé des distances de Cook	groupées non groupées	identification des observations ayant une influence sur l'estimation des paramètres.
graphique indexé des statistiques L	groupées non groupées	identification des observations ayant une influence sur le changement de déviance entre 2 modèles emboîtés.