**Logistic Regression**

Une image contenant texte, ligne, capture d’écran, Tracé

Description générée automatiquement

**Régression linéaire SIMPLE de TookAction sur Age**

Une image contenant texte, capture d’écran, diagramme, Police

Description générée automatiquementUne image contenant texte, capture d’écran, diagramme, ligne

Description générée automatiquementOn remarque que les points d’observation sont très loin des valeurs prédites par la droite. Alors La droite de régression linéaire n’est pas le modèle adéquate pour prédire l’action TookAction avec Age.

Une image contenant texte, Police, logo, vert

Description générée automatiquement

Avec p : la probabilité que l’évènement se réalise.

**Une image contenant texte, capture d’écran, diagramme

Description générée automatiquementLogistic Regression of TookAction sur Age**

Model 1: Logit, using observations 1-100

Dependent variable: TookAction

Standard errors based on Hessian

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | *Coefficient* | *Std. Error* | *z* | *p-value* |  |
| const | −22.5768 | 4.79476 | −4.709 | <0.0001 | \*\*\* |
| Age | 0.557366 | 0.118128 | 4.718 | <0.0001 | \*\*\* |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Mean dependent var | 0.400000 |  | S.D. dependent var | 0.492366 |
| McFadden R-squared | 0.652483 |  | Adjusted R-squared | 0.622766 |
| Log-likelihood | −23.38827 |  | Akaike criterion | 50.77655 |
| Schwarz criterion | 55.98689 |  | Hannan-Quinn | 52.88526 |

Number of cases 'correctly predicted' = 90 (90.0%)

f(beta'x) at mean of independent vars = 0.492

Likelihood ratio test: Chi-square(1) = 87.8258 [0.0000]

Predicted

0 1

Actual 0 55 5

1 5 35

**Logistic Regression of TookAction sur Age et Gender(Male)**

On a dummify la variable Gender en gardant Male.

Une image contenant capture d’écran, texte, diagramme

Description générée automatiquement

Model 2: Logit, using observations 1-100

Dependent variable: TookAction

Standard errors based on Hessian

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | *Coefficient* | *Std. Error* | *z* | *p-value* |  |
| const | −38.1520 | 9.98690 | −3.820 | 0.0001 | \*\*\* |
| Age | 0.887226 | 0.231787 | 3.828 | 0.0001 | \*\*\* |
| Male | 4.43738 | 1.49193 | 2.974 | 0.0029 | \*\*\* |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Mean dependent var | 0.400000 |  | S.D. dependent var | 0.492366 |
| McFadden R-squared | 0.779993 |  | Adjusted R-squared | 0.735417 |
| Log-likelihood | −14.80675 |  | Akaike criterion | 35.61349 |
| Schwarz criterion | 43.42901 |  | Hannan-Quinn | 38.77657 |

Number of cases 'correctly predicted' = 94 (94.0%)

f(beta'x) at mean of independent vars = 0.492

Likelihood ratio test: Chi-square(2) = 104.989 [0.0000]

Predicted

0 1

Actual 0 58 2

1 4 36

Voir la table de prédiction dans tab\_predict.pdf

* False Positive(FP : Erreur de 1ere espèce), c’est l’erreur commise par le modèle lorsqu’il predit que qq va se passer alors qu’en réalité NON.
* False Negative(FN : Erreur de 2e espèce), c’est l’erreur commise par le modèle lorsqu’il prédit que qq ne va pas se passer alors qu’en réalité OUI.
* Une image contenant texte, capture d’écran, diagramme, nombre

  Description générée automatiquementMatrice de confusion

**Interprétations des coefficients de la régression logistique**

Prenons la Logistic Regression of TookAction sur Age et Gender(Female), on a:

Model 3: Logit, using observations 1-100

Dependent variable: TookAction

Standard errors based on Hessian

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | *Coefficient* | *Std. Error* | *z* | *p-value* |  |
| const | −33.7146 | 8.83304 | −3.817 | 0.0001 | \*\*\* |
| Age | 0.887226 | 0.231787 | 3.828 | 0.0001 | \*\*\* |
| Female | −4.43738 | 1.49193 | −2.974 | 0.0029 | \*\*\* |

* Age >0,i.e. si vous prenez 2 femmes, alors la femme la plus âgé a plus de chance de cliquer sur le lien de l’offre que la femme la moins âgé.
* Female < 0, i.e.si on maintient toutes les autres variables indépendantes constantes, i.e. si on prend 2 personnes du même âge, un homme et une femme du même âge alors la femme a moins de cliquer sur l’offre que l’homme.

Une image contenant texte, capture d’écran, Police, nombre

Description générée automatiquement