

- **Quel est l'intérêt de la régression logistique ?**

Comme pour la régression linéaire et le modèle de Cox, le but de la régression logistique est de caractériser les relations entre une variable dépendante (ou variable à expliquer) et une seule (régression logistique simple) ou plusieurs variables prises en compte simultanément (régression logistique multiple). Il s'agit donc d'un modèle permettant de relier la variable dépendante (Y) à des variables explicatives ($X_1, X_2, X_3, \dots, X_n$).

À la différence de la régression linéaire (où la variable à expliquer est une variable quantitative) et du modèle de Cox (où la variable à expliquer est une variable censurée), la régression logistique s'applique lorsque la variable à expliquer (Y) est qualitative. Dans la recherche biomédicale et plus particulièrement en ophtalmologie, il est très fréquent de rencontrer de telles variables, le plus souvent de type binaire comme par exemple la présence (ou l'absence) d'une maladie, d'une récurrence ou d'une complication. Les variables explicatives (X_i) peuvent être, quant à elles, qualitatives ou quantitatives. Ces variables indépendantes sont susceptibles d'influencer la survenue ou non de la maladie, de la récurrence ou de la complication.

- **Donner des exemples de problèmes auxquels la régression logistique permet de**

Répondre ?

La régression logistique est l'un des modèles d'analyse multivariée les plus couramment utilisés en épidémiologie.

Dans le domaine des assurances, elle permet de cibler une fraction de la clientèle qui sera sensible à une police d'assurance sur tel ou tel risque particulier.

Dans le domaine bancaire, pour détecter les groupes à risque lors de la souscription d'un crédit.

En économétrie, pour expliquer une variable discrète. Par exemple, les intentions de vote aux élections.

- **Combien y a-t-il de types de régressions logistiques ? Lesquels ?**

Modèle logit pour données binaires

Modèle logit pour données nominales

Modèle logit basé sur les probabilités cumulées

Modèle logit basé sur les catégories adjacentes supérieures cumulées

Modèle basé sur les logits des catégories adjacentes

· **Pourquoi le nom de « régression logistique » ?**

La régression logistique est l'un des algorithmes de base et populaires pour résoudre un problème de classification.

Il est nommé «régression logistique», car sa technique sous-jacente est assez similaire à la régression linéaire.

Le terme «logistique» est emprunté à la fonction Logit utilisée dans cette méthode de classification

· **Que sont les False Positives et les False Negatives ?**

Un faux positives est une valeur classifieur positivement alors qu'elle est négative

Un faux negatives est une valeur classifieur négativement alors qu'elle est positif

l'ensemble représente les erreurs de notre model

· **Qu'appelle-t-on matrice de confusion ? Quel autre nom a-t-elle ?**

une matrice qui mesure la qualité d'un système de classification. (grace aux faux positif et aux faux négatif)

une matrice d'erreur

· **Que sont les notions de :**

◦ *performance ou justesse ou accuracy,*

Cela représente la performance global du model

◦ *d'erreur,*

Cela représente le taux d'erreur du model

- *de précision ou valeur prédictive positive,*
la probabilité que la condition soit présente lorsque le test est positif
- *de sensibilité ou rappel ou recall,*
la probabilité que la condition ne soit pas présente lorsque le test est négatif
- *de spécificité ?*
la sensibilité (ou sélectivité) d'un test mesure sa capacité à donner un résultat positif lorsqu'une hypothèse est vérifiée.

· **Comment interprète-t-on les coefficients de la régression logistique ?**

Cela permet de connaître l'importance d'une variable explicative x pour l'explication d'une variable expliquée x

· **Que sont les odds ratio (rapport de cotes) ?**

le degré de dépendance entre des variables aléatoires qualitatives

· **Quel est le lien entre les odds ratio et les coefficients ?**

·

Table des matières

• Quel est l'intérêt de la régression logistique ?	1
• Donner des exemples de problèmes auxquels la régression logistique permet de	1
• Combien y a-t-il de types de régressions logistiques ? Lesquels ?	1
<i>Modèle logit pour données binaires</i>	<i>1</i>
<i>Modèle logit pour données nominales</i>	<i>2</i>
<i>Modèle logit basé sur les probabilités cumulées</i>	<i>2</i>
<i>Modèle logit basé sur les catégories adjacentes supérieures cumulées</i>	<i>2</i>
<i>Modèle basé sur les logits des catégories adjacentes</i>	<i>2</i>