

Étude cas-témoin sur l'infarctus du myocarde

Projet de biostatistiques

Biostatistiques

LANCON Cindy, FRISON Victor

Encadrant : Davit Varron

2026-01-15

Plan

1 Introduction

► Introduction

► Regression logistique simple / univarié

Introduction

1 Introduction

Design et Population :

- **Type** : Étude multicentrique sur l'infarctus du myocarde aigu.
- **Echantillon global** : $N = 520$ patients.

Constitution des groupes :

- **260 Cas** : Patients hospitalisés pour un 1^{er} infarctus confirmé.
- **260 Témoins** : Patients hospitalisés (mêmes centres) pour pathologies non cardio-vasculaires, sans antécédents coronariens.

Variables et Modélisation

Variable réponse (Y) :

Statut du patient

Modalité de référence : Témoin

Objectif : Modéliser la probabilité d'être Cas.

Facteurs de risque étudiés (X) :

Cliniques : Âge, Sexe, IMC, HTA, Diabète, Cholestérol.

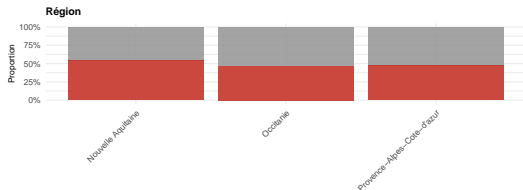
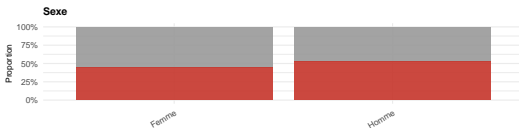
Mode de vie : Activité physique, Alcool, Région.

Tabagisme : Jamais / Ancien / Actuel.

Réf Tabac : Jamais fumeur.

Analyse exploratoire et Statistiques Descriptives

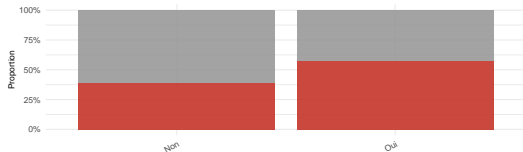
1 Introduction



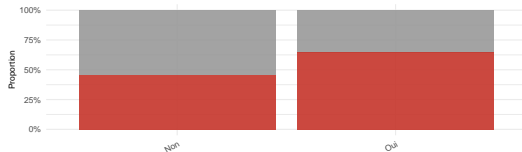
Analyse exploratoire et Statistiques Descriptives

1 Introduction

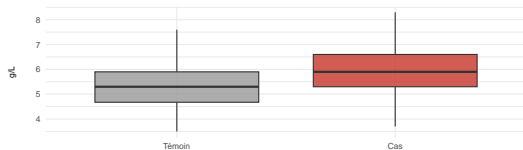
Hypertension



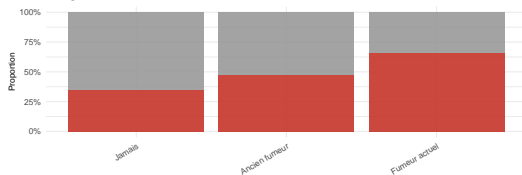
Diabète



Cholestérol Total

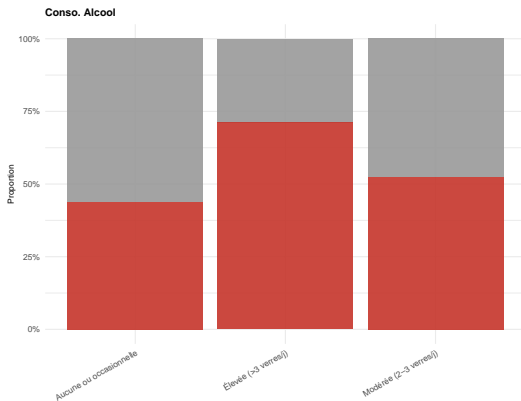
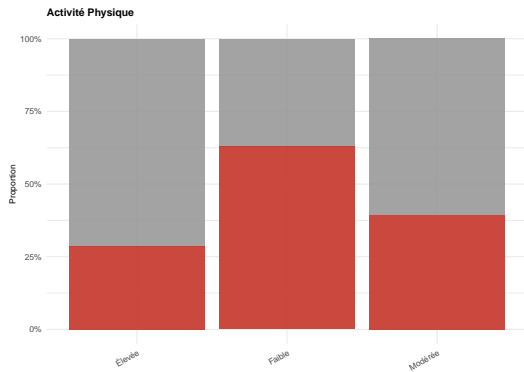


Tabagisme



Analyse exploratoire et Statistiques Descriptives

1 Introduction



Plan

2 Regression logistique simple / univarié

► Introduction

► Regression logistique simple / univarié

Variable Âge

2 Regression logistique simple / univarié

Résultats du modèle logistique :

- Coefficient estimé ($\hat{\beta}_{age}$) : 0.083
- Odds Ratio (OR) : 1.09
- Intervalle de Confiance à 95% : [1.06 ; 1.11]
- Significativité (p-value) : $p < 0.001$

Analyse de l'Odds Ratio : Pour chaque année d'âge supplémentaire, l'odds de faire un infarctus est multiplié par environ **1.09**. Concrètement, le risque augmente de **6% à 11%** par an.

Conclusion

L'intervalle de confiance ne contient pas la valeur 1 et la p-value est extrêmement faible. **L'âge est donc un facteur de risque significativement associé à la survenue de l'infarctus.**

Résultats du modèle logistique :

- Coefficient estimé ($\hat{\beta}_{imc}$) : 0.148
- Odds Ratio (OR) : 1.16
- Intervalle de Confiance à 95% : [1.10 ; 1.22]
- Significativité (p-value) : $p < 0.001$

Analyse de l'Odds Ratio : Pour chaque unité d'IMC supplémentaire, l'odds d'être malade est multiplié par environ **1.16**. Concrètement, le risque augmente de **10% à 22%** par unité.

Conclusion

L'intervalle ne contient pas 1 et la p-value est extrêmement faible (6.7×10^{-10}). **L'IMC est donc significativement associé à la survenue de l'infarctus.**

Résultats du modèle logistique :

- Coefficient estimé ($\hat{\beta}_{chol}$) : 0.785
- Odds Ratio (OR) : 2.19
- Intervalle de Confiance à 95% : [1.79 ; 2.71]
- Significativité (p-value) : $p < 0.001$

Analyse de l'Odds Ratio : Pour chaque unité de cholestérol supplémentaire, l'odds est multiplié par environ **2.19**. Le risque augmente considérablement, de **79% à 171%** par unité.

Conclusion

L'intervalle est nettement supérieur à 1. **Le cholestérol total est un facteur de risque majeur et hautement significatif.**

Variable Sexe

2 Regression logistique simple / univarié

Résultats du modèle logistique (Réf : Femme) :

- Coefficient estimé ($\hat{\beta}_{Homme}$) : 0.307
- Intervalle de Confiance de l'OR à 95% : [0.95 ; 1.95]
- Significativité (p-value) : > 0.05 (Non significatif)

Analyse : L'intervalle de confiance de l'Odds Ratio contient la valeur 1 ([0.95; 1.95]). Cela signifie qu'il n'y a pas de différence de risque statistiquement prouvée entre les hommes et les femmes dans cet échantillon.

Conclusion

La p-value est supérieure au seuil de 5% et l'IC contient 1. **Le sexe n'est pas statistiquement associé à la survenue de l'infarctus.**

Variable Tabagisme

2 Regression logistique simple / univarié

Résultats par rapport aux non-fumeurs (Réf : Jamais) :

- **Ancien Fumeur :**
 - OR IC 95% : [1.07 ; 2.61] ($p = 0.023$)
- **Fumeur Actuel :**
 - OR IC 95% : [2.34 ; 5.56] ($p < 0.001$)

Analyse des risques :

- Être **ancien fumeur** augmente le risque (OR min : 1.07).
- Être **fumeur actuel** l'augmente drastiquement (OR entre 2.34 et 5.56).

Conclusion

Le test global est très significatif ($p < 0.001$). **Le tabagisme est fortement associé à la maladie, avec un effet dose-réponse (Actuel > Ancien > Jamais).**

Variable Hypertension

2 Regression logistique simple / univarié

Résultats du modèle logistique (Réf : Non) :

- Coefficient estimé ($\hat{\beta}_{HTA}$) : 0.75
- Intervalle de Confiance de l'OR à 95% : [1.47 ; 3.03]
- Significativité (p-value) : $p < 0.001$ (3.6×10^{-5})

Analyse de l'Odds Ratio : Le fait d'être hypertendu multiplie le risque (l'odds) d'être un cas par au moins **1.47** et au plus **3.03** par rapport aux sujets sains.

Conclusion

L'intervalle de confiance est nettement supérieur à 1. L'hypertension artérielle est un facteur de risque majeur statistiquement associé à l'infarctus.

Variable Diabète

2 Regression logistique simple / univarié

Résultats du modèle logistique (Réf : Non) :

- Coefficient estimé ($\hat{\beta}_{diabète}$) : 0.77
- Odds Ratio (OR) : 2.16 ($e^{0.77}$)
- Intervalle de Confiance à 95% : [1.42 ; 3.34]
- Significativité (p-value) : $p < 0.001$ (3.2×10^{-4})

Analyse de l'Odds Ratio : Être diabétique multiplie l'odds d'être un cas par au moins **1.42** et au plus **3.34** par rapport à un sujet non diabétique.

Conclusion

La p-value est très faible et l'intervalle exclut 1. Le **diabète est statistiquement associé à la contraction de la maladie.**

Résultats par rapport à une activité Élevée (Réf) :

- **Activité Faible :**
 - OR IC 95% : [2.48 ; 7.51] ($p < 0.001$)
- **Activité Modérée :**
 - OR IC 95% : [0.91 ; 2.93] ($p = 0.104$, Non Sign.)

Analyse des risques : Une activité physique **faible** augmente drastiquement le risque ($OR > 2.48$). En revanche, la différence entre activité modérée et élevée n'est pas significative.

Conclusion

Le test global est significatif ($p < 0.001$). **La sédentarité (activité faible) est un facteur de risque majeur par rapport à une activité sportive régulière.**

Variable Alcool

2 Regression logistique simple / univarié

Résultats par rapport à Aucune/Occas. (Réf) :

- **Consommation Élevée :**
 - OR IC 95% : [1.80 ; 5.79] ($p < 0.001$)
- **Consommation Modérée :**
 - OR IC 95% : [0.96 ; 2.08] ($p = 0.08$, Non Sign.)

Analyse des risques : Une consommation **élevée** multiplie le risque par environ 3 (OR estimé à 3.15). Une consommation modérée ne montre pas de sur-risque significatif.

Conclusion

Le test global est significatif ($p < 0.001$). **L'abus d'alcool (consommation élevée) est fortement associé à la maladie.**