

# Examen de Biostatistique

3<sup>ème</sup> année, Groupes C et D  
Année universitaire 2024-2025, Période 2

Enseignante : Rym Jaroudi

Ecole Supérieure de la Statistique et de l'Analyse de l'Information  
Université de Carthage

Samedi 11 janvier 2025  
9h - 10h30

## Instructions

L'examen comprend 3 parties :

- Partie 1 : Questions théoriques (Questions 1 à 5) – 8 points
- Partie 2 : Interprétation et calculs (Questions 6 et 7) – 8 points
- Partie 3 : QCM (Questions 8 à 10) – 4 points

L'examen est noté sur 20 points. Le barème est indiqué pour chaque section. Répondez à toutes les questions. **Aucun document n'est autorisé pendant l'examen.**

## 1 Partie 1 : Questions théoriques (8 points)

**Question 1 :** Expliquez les concepts suivants : (2 points)

- a) In silico
- b) Incidence

**Question 2 :** Associez chaque exemple de donnée au type approprié (quantitative continue, quantitative discrète, qualitative nominale, qualitative ordinale): (2 points)

- |  |   |
|--|---|
| a) Nombre d'hospitalisations par an                          | c) Taux de glucose sanguin (en g/L)                         |
| b) Classe de gravité d'une maladie (faible, modérée, élevée) | d) Niveau d'éducation (primaire, secondaire, universitaire) |

**Question 3 :** Expliquez brièvement les concepts in vivo et in vitro en mettant en avant leurs différences principales. (2 points)

**Question 4 :** Le gène BRCA1 est impliqué dans la réparation de l'ADN et est associé à un risque accru de cancer du sein et d'ovaire lorsqu'il est muté. Une étude examine si l'expression de BRCA1 diffère entre les tissus tumoraux et les tissus sains. Formulez l'hypothèse nulle ( $H_0$ ) et l'hypothèse alternative ( $H_1$ ) pour cette situation. (1 point)

**Question 5 :** Donnez un exemple d'erreur courante en biostatistique et expliquez comment l'éviter. (1 point)

## 2 Partie 2 : Interprétation et calculs (8 points)

**Question 6 :** Un modèle de régression logistique a été ajusté pour prédire la probabilité qu'un patient développe une maladie cardiovasculaire ( $Y = 1$ ) en fonction de :

- Âge ( $X_1$ ) : en années
- IMC ( $X_2$ ) : indice de masse corporelle

Les résultats obtenus sont :

Variable	Coefficient ( $\beta$ )	Odds Ratio (OR)	p-valeur
Âge	0.06	1.06	0.01
IMC	0.2	1.22	0.03

- Interprétez l'effet de l'âge ( $OR = 1.06$ ) et de l'IMC ( $OR = 1.22$ ) sur le développement de la maladie cardiovasculaire, indiquez si ces effets sont statistiquement significatifs. (1 point)
- Peut-on dire que l'effet de l'âge sur le développement de la maladie cardiovasculaire est cliniquement significatif ? (1 point)
- Pourquoi utilise-t-on une régression logistique dans ce cas et non une régression linéaire ? (1 point)
- Calculez la probabilité qu'un individu âgé de 50 ans avec un IMC de 30 développe une maladie cardiovasculaire si l'intercept du modèle est  $-3.5$ . (1 point)

**Question 7 :** Un chercheur souhaite étudier l'association entre un facteur  $F+$  (consommation élevée de sel) et une maladie  $M+$  (hypertension). Voici les données :

- 500 personnes consomment beaucoup de sel, dont 200 ont une hypertension.
  - 800 personnes consomment peu de sel, dont 240 ont une hypertension.
- Calculez la prévalence de l'hypertension dans la population totale. (1 point)
  - Calculez le risque relatif (RR) pour les consommateurs élevés de sel par rapport aux faibles consommateurs. Interprétez le résultat obtenu. (2 points)
  - Quel biais pourrait affecter cette étude et comment pourrait-il influencer les résultats ? (1 point)



### 3 Partie 3 : Questions à choix multiple (4 points)

**Question 8 :** Une étude de survie vise à analyser le temps jusqu'à la récurrence d'un cancer chez des patients répartis en deux groupes :

- Groupe A : traitement standard
- Groupe B : traitement expérimental

Les résultats du modèle de Cox ajusté sont les suivants :

Variable	Coefficient ( $\beta$ )	Hazard Ratio (HR)	p-valeur
Groupe B (vs A)	-0.7	0.50	0.04
Âge	0.03	1.03	0.10

Sélectionnez la réponse correcte pour chaque question.

1. La p-valeur pour le groupe B est de 0.04. Que peut-on conclure ? (1 point)
  - a) L'effet du traitement expérimental est significatif au seuil de 5 %.
  - b) L'effet du traitement expérimental est significatif au seuil de 1 %.
  - c) L'effet du traitement expérimental n'est pas significatif.
  - d) Le traitement expérimental augmente le risque de récurrence.
2. Que signifie le Hazard Ratio (HR) de 1.03 pour la variable Âge ? (1 point)
  - a) À chaque année supplémentaire, le risque instantané de récurrence diminue de 3 %.
  - b) À chaque année supplémentaire, le risque instantané de récurrence augmente de 3 %.
  - c) Le risque instantané de récurrence est multiplié par 7 chaque année.
  - d) L'âge n'a aucun effet sur le risque instantané de récurrence.

**Question 9 :** Dans la réglementation des essais cliniques, quel logiciel est le plus couramment utilisé pour le traitement des données issues des essais cliniques, notamment pour la gestion des bases de données ? Sélectionnez la réponse correcte. (1 point)

- a) SAS
- b) R
- c) Python

**Question 10 :** Quelle condition doit être vérifiée pour effectuer un test ANOVA valide ? Sélectionnez la réponse correcte. (1 point)

- a) Les données doivent être indépendantes, mais la normalité et l'homogénéité des variances ne sont pas nécessaires.
- b) Les données doivent suivre une distribution normale et les variances doivent être homogènes.
- c) Il n'y a pas de condition à vérifier pour effectuer un test ANOVA.