

## **QCM - Introduction à l'inférence causale**

**Instructions : Pour chaque question, choisissez la bonne réponse.**

1. **Qu'est-ce que l'inférence causale ?**
  - a) Une méthode pour établir des relations de corrélation
  - b) Un processus visant à identifier un lien de causalité entre des variables
  - c) Une analyse purement descriptive des données
  - d) Une technique de prédiction
2. **Quelle est la principale difficulté de l'inférence causale ?**
  - a) La collecte de données
  - b) L'établissement de relations de corrélation
  - c) L'impossibilité d'observer directement les contrefactuels
  - d) L'utilisation de modèles statistiques complexes
3. **Qu'est-ce qu'un contrefactuel ?**
  - a) Un événement observé dans les données
  - b) Un événement hypothétique qui aurait pu se produire sous une condition différente
  - c) Une variable explicative
  - d) Une hypothèse confirmée
4. **Quel modèle est souvent utilisé pour formuler l'inférence causale ?**
  - a) Le modèle de régression linéaire
  - b) Le modèle de variables latentes
  - c) Le cadre de Rubin (Rubin Causal Model)
  - d) Le modèle des séries temporelles
5. **Quelle est la définition d'un traitement dans une étude causale ?**
  - a) Un groupe de contrôle
  - b) Une variable indépendante
  - c) Une intervention ou une exposition dont on veut évaluer l'effet
  - d) Un facteur aléatoire
6. **Pourquoi utilise-t-on des expériences randomisées contrôlées (ERC) en inférence causale ?**
  - a) Pour observer toutes les variables possibles
  - b) Pour établir des relations de corrélation
  - c) Pour garantir une attribution aléatoire et éviter les biais de sélection
  - d) Parce qu'elles sont plus simples à analyser
7. **Quelle méthode est utilisée lorsque les ERC ne sont pas possibles ?**
  - a) La régression linéaire
  - b) Les méthodes quasi-expérimentales
  - c) L'analyse descriptive
  - d) La classification

8. **Quelle méthode quasi-expérimentale repose sur une hypothèse de quasi-randomisation ?**
- a) La régression sur variables instrumentales
  - b) La méthode des doubles différences (DID)
  - c) La régression logistique
  - d) L'apprentissage supervisé
9. **Quel est le rôle des variables instrumentales en inférence causale ?**
- a) Remplacer les variables indépendantes
  - b) Aider à identifier les effets causaux en éliminant les biais de confusion
  - c) Tester la robustesse d'un modèle
  - d) Déterminer les corrélations
10. **Pourquoi les modèles de régression simples ne suffisent-ils pas toujours pour établir la causalité ?**
- a) Parce qu'ils ne permettent pas de mesurer les contrefactuels
  - b) Parce qu'ils sont trop complexes
  - c) Parce qu'ils exigent trop de données
  - d) Parce qu'ils ne permettent pas d'analyser les séries temporelles
11. **Quel est l'objectif principal des modèles de régression sur discontinuité (RD) ?**
- a) Tester des effets à long terme
  - b) Évaluer l'effet causal d'une intervention en exploitant un seuil arbitraire
  - c) Identifier les facteurs prédictifs d'une variable
  - d) Comparer des groupes homogènes
12. **Que signifie "DID" en inférence causale ?**
- a) Différence en différences
  - b) Distribution inverse des données
  - c) Dépendance interne des données
  - d) Décomposition informative des données
13. **Quels sont les principaux problèmes liés à l'omission d'une variable confondante ?**
- a) Biais dans l'estimation des effets causaux
  - b) Augmentation du bruit dans les données
  - c) Réduction du nombre d'observations
  - d) Amélioration de la précision du modèle
14. **Quel outil graphique est souvent utilisé pour représenter des relations causales ?**
- a) Les diagrammes de Venn
  - b) Les diagrammes de dispersion
  - c) Les graphes causaux dirigés (DAGs)
  - d) Les arbres de décision

15. **Quel principe clé est à la base de la méthode des doubles différences ?**
- a) Comparer l'évolution d'un groupe traité et d'un groupe de contrôle avant et après une intervention
  - b) Appliquer une double randomisation
  - c) Utiliser deux variables instrumentales
  - d) Éliminer toutes les variables indépendantes
16. **Quel est l'intérêt des méthodes de machine learning pour l'inférence causale ?**
- a) Elles permettent d'identifier des corrélations fortes
  - b) Elles automatisent le processus d'analyse causale
  - c) Elles aident à mieux contrôler les variables confondantes et les interactions complexes
  - d) Elles remplacent les expériences randomisées
17. **Quel concept permet d'éviter le biais de sélection dans une étude causale ?**
- a) L'hétéroscédasticité
  - b) L'attribution aléatoire des traitements
  - c) L'analyse en composantes principales
  - d) L'optimisation des hyperparamètres
18. **Comment interpréter un effet causal estimé avec une approche quasi-expérimentale ?**
- a) Comme une approximation de l'effet réel, sous certaines hypothèses
  - b) Comme une vérité absolue
  - c) Comme une simple corrélation
  - d) Comme un test de significativité
19. **Pourquoi est-il important d'examiner la validité des hypothèses en inférence causale ?**
- a) Pour confirmer que la régression linéaire est bien appliquée
  - b) Pour garantir la robustesse des conclusions causales
  - c) Pour vérifier l'efficacité du modèle prédictif
  - d) Pour obtenir un meilleur ajustement des données
20. **Quel auteur est souvent cité comme une référence en inférence causale ?**
- a) Karl Pearson
  - b) Judea Pearl
  - c) Francis Galton
  - d) Ronald Fisher