

M2 : EES – Etudes Economiques & Statistiques **Examen Data Science, novembre 2025, Durée 1h 30**

Partie 1 : Questions à Choix Multiple – Test d'évaluation des connaissances

Barème partie 1 : (24 questions - 6 points : Bonne réponse = + 0.25 pt, Pas de réponse = 0 pt, Mauvaise réponse = - 0.25 pt)

Répondre par **Vrai** ou **Faux** (entourer la bonne réponse) - Chaque affirmation peut être Vraie ou Fausse. Elle est considérée comme vraie si elle est TOUJOURS vraie. Par contre si l'affirmation est absurde ou si elle n'a pas de sens, elle est considérée comme fausse. Il peut y avoir plusieurs affirmations vraies ainsi que plusieurs affirmations fausses dans chaque groupe d'affirmations.

1) Analyse en Composantes Principales (ACP) normée d'un tableau de données X à n individus et p variables :

- a) la représentation graphique des individus est un cercle de corrélation. F
- b) l'ACP est une méthode de réduction de dimension. V
- c) les coordonnées des points variables sur le plan factoriel sont ceux des corrélations (variables / facteurs). V
- d) les variables illustratives peuvent être mixtes. V

2) Analyse Factorielle des Correspondances (AFC) - Test d'indépendance du chi-deux sur une table de contingence croisant deux variables qualitatives respectivement à p et à q modalités :

- a) l'AFC consiste à localiser les éventuelles corrélations entre les modalités des deux variables qualitatives. F
- b) l'AFC est une double ACP des profils lignes et colonnes. V
- c) l'hypothèse nulle du test est H_0 : les deux variables sont dépendantes. F
- d) plus la p-value est grande plus les variables sont indépendantes. V

3) Analyse des correspondances Multiples (ACM) de p variables qualitatives totalisant q modalités :

- a) L'ACM est une généralisation de l'ACP. F
- b) le nombre de composantes principales de l'ACM est égal à $q \times p$. F
- c) le tableau de Burt est la juxtaposition de tableaux de contingences. V
- d) les facteurs ou composantes principales de l'ACM sont des variables nominales. F

4) Analyse Factrielle Discriminante (AFD) à p variables explicatives et une variable à expliquer à q modalités ($p > q$) :

- a) l'AFD consiste à minimiser l'inertie ou l'information Intra-classes. V
- b) L'AFD est une méthode de modélisation prédictive du machine learning au service de la Data Science pour la modélisation des données. V
- c) l'AFD sur variables explicatives qualitatives est une AFD sur les facteurs principaux de l'ACP des variables explicatives. F
- d) Permet de séparer et de différencier au mieux les q classes de la variable qualitative à expliquer. V

5) Classifications : Nuées dynamiques et Classification Hiérarchique Ascendante (CHA)

- a) dans le cas d'une CHA, le nombre de classes n'est pas fixé par l'utilisateur. V
- b) La méthode des nuées dynamique est une méthode hiérarchique. F
- c) ces techniques sont des classifications dites non supervisées. V
- d) Lorsque la taille de l'échantillon est grande on applique la méthode des nuées dynamique. V

6) Modélisation prédictive :

- a) Les ANOVA et ANCOVA sont des modèles décisionnels de classement. F
- b) L'ANOVA est un modèle de comparaison de moyennes. V
- c) l'ANCOVA est un modèle qui permet d'analyser l'effet de variables explicatives qualitatives sur une variable continue à expliquer. F
- d) Les distributions des populations considérées dans l'ANOVA doivent suivre des lois normales. V

Compléter, signer et retourner ce QCM avec votre copie d'examen.

Nom	Prénom	Signature