

Ce travail est à réaliser en équipe (minimum deux, maximum quatre personnes). Vous devez remettre votre rapport au format PDF ainsi que votre code **R** (ou un fichier Quarto).

Utilisez la convention de nomenclature `d2_matricule.extension`, où `matricule` est le matricule de l'étudiant(e) qui soumet le rapport et `extension` est un de `pdf`, `qmd`, `R`.

1. Plusieurs systèmes de technologies de l'information servent à centraliser les requêtes pour des services. On considère le temps d'attente (en heures, variable *duree*) avant la résolution d'un billet, jusqu'à concurrence de 120 heures.
  - (a) Une des opérations préliminaires à l'analyse de survie est le calcul du temps d'attente sur la base des registres informatiques; les données sont typiquement au format `YYYY:MM:DD HH.MM`. Si l'entreprise opère uniquement de 9h à 17h du lundi au vendredi, décrivez brièvement comment vous pourriez prendre en compte cette particularité lors du calcul du temps entre l'ouverture et la fermeture du billet de support si vous aviez à calculer *duree* vous-même.
  - (b) Estimez la fonction de survie séparément pour chaque type de billets à l'aide de l'estimateur de Kaplan–Meier.
    - i. Rapportez les quartiles de la survie pour les requêtes.
    - ii. Rapportez la probabilité estimée qu'un billet pour un problème soit encore ouverte après 8h (soit un jour ouvrable de traitement).
    - iii. Expliquez dans vos mots pourquoi certains quartiles pour les “problèmes” ne sont pas disponibles (valeurs manquantes).
    - iv. Produisez un graphique des deux courbes de survie estimées avec des intervalles de confiance à niveau 95%. Est-ce que les courbes sont significativement différentes l'une de l'autre?
  - (c) Ajustez un modèle de Cox pour prendre en compte l'effet du type de demande, du niveau de difficulté (traité comme variable catégorielle initialement) et des commentaires.
    - i. On suppose que les commentaires reflètent des allers-retours entre la personne qui fait une demande de service et le personnel TI et entraîne des temps de traitement plus longs en moyenne. Est-ce que les données corroborent cette hypothèse?
    - ii. Toute chose étant égale par ailleurs, pour quel niveau de difficulté est-ce que le temps de survie est le plus long? et le plus court? Justifiez votre réponse.
    - iii. Interprétez le coefficient du modèle de Cox associé au nombre de commentaires en terme de pourcentage d'augmentation ou de diminution du risque.
    - iv. Vérifiez le postulat de risques proportionnels pour le modèle qui inclut *niveau* comme variable catégorielle.
    - v. Ajustez le modèle en stratifiant par type de demande et rapportez les coefficients estimés dans un tableau.
    - vi. On pourrait traiter le *niveau* comme variable continue en supposant une gradation constante entre les niveaux de difficulté. Est-ce que les données supportent cette hypothèse? Rapportez la statistique du test de rapport de vraisemblance, la valeur-*p* et la conclusion de votre test d'hypothèse. *Indication : comparez les deux modèles avec `anova`*

2. Les données de cette question, extraites de ResearchBox 159 et mises à disposition sous licence CC BY 4.0, ont servi à construire une échelle psychométrique pour mesurer l'avarice. Les données de  $n = 300$  participant(e)s ont été collectées à partir de Amazon M'Turk.

Zeelenberg M, Seuntjens TG, van de Ven N, Breugelmans SM (2021). *Dispositional Greed Scales*. European Journal of Psychological Assessment. 1–10. doi : 10.1027/1015-5759/a000647

Les participant(e)s devaient répondre à 30 questions'. Les instructions étaient : «Indiquez pour chaque énoncé à quel point vous êtes en accord» à l'aide d'une échelle de Likert avec comme choix «fortement en désaccord» (1), «en désaccord» (2), «ni en accord, ni en désaccord» (3), «d'accord» (4) et «fortement d'accord». Les questions sont reproduites ci-dessous.

- (a) Est-il nécessaire de faire des modifications aux données brutes avant l'analyse (standardisation, encodage inverse des échelles)? Justifiez votre réponse.
- (b) Effectuez une analyse factorielle exploratoire à l'aide de la méthode des composantes principales avec rotation varimax (option par défaut).
  - i. À l'aide des critères couverts en classe (valeurs propres de Kaiser, coude), déterminez un nombre adéquat de facteurs à employer. Justifiez adéquatement votre choix.
  - ii. Rapportez le pourcentage de la variance expliquée par votre méthode.
  - iii. Étiquetez les facteurs obtenus et expliquez brièvement leur signification.
  - iv. En utilisant les facteurs choisis, créez des échelles et vérifiez leur cohérence interne via le  $\alpha$  de Cronbach.

### Liste des questions

- Q1. I like to collect expensive things
- Q2. At work/school, I keep good ideas to myself so that only I can get credit for them in the long run
- Q3. No matter how much I have of something, I always want more
- Q4. My actions are strongly focused on material things
- Q5. I get the most fun out of buying myself all sorts of things
- Q6. A simple basic life is sufficient for me
- Q7. When I think about all the things I have, my first thought is about what I would like to have next
- Q8. One can never have enough
- Q9. When something is being shared, I try to get as big a share as possible
- Q10. One can never have too much money
- Q11. I enjoy being a part of exclusive clubs or groups that are not open to everyone
- Q12. I believe that money is essential; friends are replaceable
- Q13. I consider myself successful if I have a job that pays a lot of money
- Q14. The pursuit of more and better is an important goal in life for me
- Q15. No matter how much I have, I always want more
- Q16. I am easily satisfied with what I've got
- Q17. "I want it all" would be a good motto for me
- Q18. Sometimes I feel a real urge to possess something

- Q19. I can't imagine having too many things
- Q20. I do not enjoy sharing positions of power
- Q21. I always want more
- Q22. Actually, I'm kind of greedy
- Q23. It doesn't matter how much I have, I'm never completely satisfied
- Q24. When I play on my own, I sometimes cheat a little
- Q25. Financially supporting the less fortunate is a priority for me
- Q26. As soon as I have acquired something, I start to think about the next thing I want
- Q27. Even when I am fulfilled, I often seek more
- Q28. My life motto is 'more is better'
- Q29. Being financially wealthy is my number one goal.
- Q30. In order to get what I want, I can accept the fact that other people may suffer damage