

Évaluation Individuelle TP N2 - Sujet B

Nom :

Prénom :

Les données analysées ici sont celles proposées dans le fichier `titanic.csv`.

Notations : Dans ce sujet, on se propose d'étudier la variable `age` selon la classe du passager (`pclass`). La classe "aisé" correspond aux passagers de `classe = 1`, celle appelée "populaire" à ceux de `classe = 2` et celle appelée "économique" à ceux de `classe = 3`. On notera X la variable âge chez la classe "aisé", Y la variable cardiaque chez les "populaire" et Z la variable cardiaque chez les "économique". On utilisera les notations suivantes :

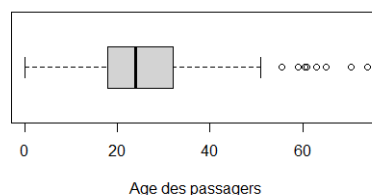
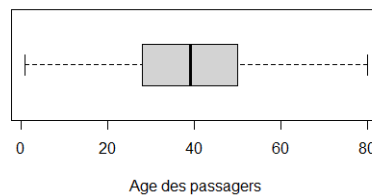
- (2 pts) Charger le jeu de données `titanic.csv` dans `tita` et extraire les échantillons de X (resp. Y et Z) qu'on affectera à `x` (resp. `y` et `z`). Indiquer les instructions utilisées, depuis le répertoire de travail contenant le fichier `titanic.csv`, pour construire `tita` et `x`.

`tita =x =`

- (3 pts : 1/2 par cellule du tableau) Compléter le tableau ci-dessous avec les estimations sans biais des espérances et écart-type des trois variables étudiées (à 10^{-3})

| | $\hat{\mu}$ | $\hat{\sigma}$ |
|-----|-------------|----------------|
| X | | |
| Y | | |
| Z | | |

- (2 pts : 0.5 première commande/1.5 deux suivantes) Indiquer les commandes permettant d'obtenir les figures ci-dessous sur une même fenêtre graphique. *Toutes les options doivent apparaître.*



.....

.....

.....

4. (2 pts) Commentaires sur le tableau et le graphique précédent :

.....

.....

5. (2 pts : 1 pt par IC) Donner des intervalles de confiance de niveaux asymptotiques (ou exacts) 99% et 90% pour le paramètre $\bar{X} = \mu_X$ (à 10^{-4}) dans le tableau suivant.

| | |
|------------------------------------|------------------------------------|
| borne inf. IC niv 99% pour μ_X | borne sup. IC niv 99% pour μ_X |
| | |
| borne inf. IC niv 90% pour μ_X | borne sup. IC niv 90% pour μ_X |
| | |

6. (2 pts) Pour une séquence de $\alpha = 0.01, 0.02, \dots, 0.3$. Représenter graphiquement (à main levée) la suite des amplitudes obtenues selon $1 - \alpha$ et indiquer approximativement à partir de quel niveau de confiance, on peut garantir une amplitude inférieure à 3 ?

.....

7. (7 pts) On donne $\mu_0 = 24$, l'âge moyen qu'un passager appartienne à la classe populaire. Faire un test statistique pour savoir si l'âge des passagers de la classe "populaire" (Y) est significativement supérieur à cette valeur de référence μ_0 . Utiliser un risque α de 5%.

- (a) (1 pt) Préciser les éventuelles conditions requises pour pouvoir faire le test mis en œuvre ainsi que la procédure utilisée.

- (b) (1.5 pts) Décrire les deux hypothèses testées :

\mathcal{H}_0 : \mathcal{H}_1 :

(c) (1,5 pts) Indiquer les instructions R exécutées pour réaliser le test précédent.

(d) (1.5 pts) Donnez la valeur de la statistique de test et la p-valeur du test :

$t_{calc} = \dots\dots\dots$ $p_{val} = \dots\dots\dots$

(e) Conclusion littérale (1,5 pts) :