

Evaluation Individuelle TP 2018-2019 : Solution Sujet 3

NOM :

GPE :

Les données utilisées dans le sujet sont dans le fichier `mtcars` et les notations décrites dans `AideExam.pdf` sont à utiliser impérativement ainsi que les notations suivantes :

Notations :

- On notera X la variable `qsec` sur la sous-population des véhicules à boîte automatique et Y sur la sous-population des véhicules à boîte manuelle.
- Les moyennes et variances de X et Y sont inconnues et seront notées resp. μ_X, μ_Y, σ_X^2 et σ_Y^2 .
- On notera Z la variable indicatrice de boîte manuelle pour les véhicules de 1973 et p son espérance.
- On sait qu'en 2018 la répartition entre boîtes manuelles et boîtes automatiques sur tous les véhicules du marché est connue et donnée par 3/4 de manuelles et 1/4 d'automatiques.

Questions :

1. (3pts) Construire les échantillons de X et Y et les affecter à `x` et `y`. Remplir le tableau suivant donnant les estimations sans biais de σ_X^2 et σ_Y^2 ainsi que les intervalles de confiance de niveau 90% (on supposera que X et Y suivent des lois normales):

Paramètre	taille éch.	e.s.b.	Borne Inf IC à 90%	Borne Sup IC à 90%
σ_X^2	13	3,2126	1,8335	7,3767
σ_Y^2	19	3,0671	1,9123	5,8791

2. (7pts) On veut savoir si en **moyenne** le type de boîte de vitesse a un effet sur `qsec`.
 - (a) Donner la ligne de commande R permettant de réaliser un graphique où figurent les boxplot des deux échantillons :
`boxplot(x,y)`
Interpréter : paramètres de centrage semblables, dispersion semblables et à noter l'indiv hors norme dans `y`
 - (b) Quel test faites vous ? test no : **5 comp. de moyennes avec échs indépendants**
 - (c) Poser les hypothèses du test :

$$\mathcal{H}_0 : \mu_X = \mu_Y \qquad \mathcal{H}_1 : \mu_X \neq \mu_Y$$

- (d) Quelles conditions doivent satisfaire X et Y pour pouvoir utiliser ce test (modèle) ?
:
 X et Y de lois normales et mêmes variances ($\sigma_X^2 = \sigma_Y^2$)
- (e) Donner la ligne de commande R permettant de réaliser le test :
`t.test(x,y,var.equal=T)`

(f) Que vaut la p-valeur du test et que décide-t-on pour $\alpha = 5\%$?

$p - val = 20,6\%$ on décide $\mathcal{H}_0 : \mu_X = \mu_Y$ car $5\% < pvalue$

(g) La conclusion suggère-t-elle de faire un test unilatéral ? Le cas échéant indiquer la commande R permettant de le réaliser et sinon justifier :

Puisqu'on a conclut $\mu_X = \mu_Y$ c'est à dire pas d'effet du type de boite sur la variable qsec on ne cherche pas le sens d'un effet

(h) Conclusion littérale du test bilatéral : **A moins de le faire avec un risque de se tromper supérieur à 20,6% on ne peut établir que le type de boite a un effet sur qsec**

3. (2pts) Le test précédent a été réalisé sous condition d'égalité des variances de X et Y . Faire le test d'égalité des variances pour vous assurer que c'est une condition raisonnablement satisfaite. Indiquer la commande R :

`var.test(x,y)`

et votre conclusion : **la pval vaut 90,1%, impossible de rejeter l'égalité à moins de risquer de se tromper à plus de 90%**

4. (4pts) On souhaite à présent savoir si la répartition de la variable type de boite de vitesse en 1973 est la même que celle bien connue de 2018.

(a) Calculer la répartition observée de **am** dans l'échantillon tiré parmi les véhicules de 1973. Compléter le tableau suivant :

modalités de am	0	1	total
freq. obs.	0,5938	0,4062	1

(b) Représenter sur un même graphique et côte à côte la répartition observée et la répartition connue pour les véhicules de 2018. Donner la commande R:

`barplot(rbind(prop.table(table(am)),c(3/4,1/4)),beside=T)`

Interpréter : **La répartition observée (en noir) semble assez différente de la répartition connue en 2018 (en gris). Il n'y a probablement pas adéquation, à vérifier par un test.**

(c) On veut vérifier l'interprétation précédemment donnée avec un test. Quel test faites-vous ? Test no : **8 : test d'adéquation du chi-deux**

(d) Compléter

\mathcal{H}_0 : la répartition de **am** en 1973 est $(3/4, 1/4)$

\mathcal{H}_1 : la répartition de **am** en 1973 est différente de $(3/4, 1/4)$

(e) Indiquer la commande R permettant de faire le test:

`chisq.test(table(am),p=c(3/4,1/4))`

(f) Donner la p-valeur : $pval = 4,12\%$ et la conclusion littérale de ce test : **On peut conclure avec un faible risque de se tromper ($>4,12\%$) que la répartition du type de boite de vitesse n'est pas la même en 1973 qu'en 2018**