

# Démarche Statistique 2020-2021

## Contrôle des connaissances (durée : 1h)

13 octobre 2020

Nom Prénom :

Cursus (L2 AA, L3 AA, M1 SAED, M1 APVV) :

Autorisé : calculatrice, une feuille de notes A4 recto-verso

### Exercice - Appréciation globale et perception sucrée

L'appréciation d'un jus d'orange par un consommateur dépend de la perception qu'il en a de sa saveur sucrée. On observe cette relation dans une expérience d'analyse sensorielle au cours de laquelle des juges, hommes et femmes, ont évalué sur une échelle allant de 1 à 10 leur appréciation d'un jus d'orange (1 : je n'aime pas, 10 : j'aime beaucoup) et leur perception de sa saveur sucrée (1 : je ne le trouve pas sucré, 10 : je le trouve très sucré).

On importe les données dans la session de travail à l'aide des commandes suivantes :

```
senso = read.table("hedo-sucree.txt",stringsAsFactors=TRUE,header=TRUE)[-1,]  
str(senso)
```

```
'data.frame': 264 obs. of 3 variables:  
 $ Sucree      : int  3 6 5 7 7 6 8 3 8 6 ...  
 $ G.Appreciation: int  3 7 3 8 7 6 9 2 9 5 ...  
 $ Sexe        : Factor w/ 2 levels "F","H": 1 1 1 1 1 1 1 1 1 1 ...
```

A l'aide de ces données, on veut répondre à la question suivante : **la relation entre l'appréciation globale d'un jus d'orange et la perception de sa saveur sucrée est-elle la même pour une femme et pour un homme ?**

### Question 1

*Dans la problématique présentée ci-dessus, quelle est la variable réponse et quelle(s) est (sont) la (les) variable(s) explicative(s) ? Donnez la nature, quantitative ou catégorielle, de ces variables (si catégorielle, quelles modalités ?).*

Réponse

## Question 2

Proposez un modèle adapté à cette problématique. Quel est le nombre de paramètres de ce modèle ?

### Réponse

Le modèle de la question 2 est estimé à partir des données disponibles. La commande suivante affiche les coefficients estimés :

```
appreciation.lm = lm(G.Appreciation~Sucree*Sexe,data=senso)
coef(appreciation.lm)

(Intercept)      Sucree      SexeH Sucree:SexeH
  1.2510516    0.6632673    1.0108254   -0.0832747
```

## Question 3

D'après les résultats ci-dessus, quelle est la valeur estimée du coefficient directeur de la droite de régression liant l'appréciation globale à la perception de la saveur sucrée pour un homme ? et pour une femme ?

### Réponse

La commande suivante calcule la somme des carrés des résidus du modèle de la question 2 :

```
RSS = sum(residuals(appreciation.lm)^2)
RSS

[1] 1055.91
```

## Question 4

Déduire du résultat précédent une estimation de l'écart-type résiduel du modèle de la question 2.

### Réponse

### Question 5

*A partir des paramètres introduits dans le modèle de la question 2, donnez l'hypothèse nulle du test de l'égalité des coefficients directeurs (ceux dont l'estimation est demandée dans la question 3) pour les hommes et les femmes.*

### Réponse

Les commandes suivantes estiment le modèle sous l'hypothèse nulle de la question 5 et calculent la somme des carrés des résidus de ce modèle :

```
appreciation.lm0 = lm(G.Appreciation~Sucree+Sexe,data=senso)
RSS0 = sum(residuals(appreciation.lm0)^2)
RSS0
```

```
[1] 1058.41
```

### Question 6

*Donnez l'expression et la valeur d'une statistique de test pour le test d'hypothèse de la question 5.*

### Réponse

### Question 7

*Quelle est la distribution sous l'hypothèse nulle de la statistique de test calculée à la question 6 ?*

### Réponse

La commande suivante calcule la probabilité qu'une variable aléatoire suivant la distribution de la question 7 soit supérieure à 1 :

```
pf(1,df1=1,df2=260,lower.tail=FALSE)
```

```
[1] 0.31824
```

### Question 8

*D'après le résultat ci-dessus, peut-on dire que les coefficients directeurs des droites de régression liant l'appréciation globale d'un jus d'orange à la perception de sa saveur sucrée pour les hommes et les femmes sont significativement différents (au seuil de 5%) ?*

### Réponse

### Question 9

*D'après les résultats suivants :*

```
appreciation.lm0 = lm(G.Appreciation~Sucree,data=senso)
appreciation.lm = lm(G.Appreciation~Sucree+Sexe,data=senso)
anova(appreciation.lm0,appreciation.lm)
```

Analysis of Variance Table

Model 1: G.Appreciation ~ Sucree

Model 2: G.Appreciation ~ Sucree + Sexe

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	262	1077				
2	261	1058	1	18.22	4.494	0.035 *

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

*quel modèle proposez-vous finalement de retenir pour la relation entre l'appréciation globale d'un jus d'orange, la perception de sa saveur sucrée et le genre du consommateur ? (justifiez)*

### Réponse