

# Démarche Statistique 2022-2023

## Contrôle des connaissances (durée : 1h)

13 octobre 2022

**Nom Prénom :**

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

**L'usage du téléphone portable est interdit, même pour sa fonction calculatrice. L'usage de l'ordinateur est interdit pendant la durée de l'épreuve.**

### Exercice - Activité cérébrale et anxiété

Une équipe de chercheurs en neurosciences a mesuré l'activité cérébrale par électro-encéphalogramme (EEG) de 15 sujets à qui des images représentant un visage étaient présentés sur un écran. Pour chaque sujet, on dispose de 8 mesures d'activité cérébrale, une pour chaque combinaison d'une image de visage neutre ou en colère, d'une durée courte ou longue d'exposition à l'image et d'une position à droite ou à gauche sur l'écran.

On donne ici un résumé des données, disponibles dans le package R permuco :

```
      score      id  visibility  emotion  direction      sex
Min.   :-0.549  S01    : 8    16ms :60  angry  :60  right:60  male  :56
1st Qu.: 5.215  S02    : 8   166ms:60  neutral:60  left :60  female:64
Median : 9.078  S03    : 8
Mean   : 9.027  S04    : 8
3rd Qu.:12.218  S05    : 8
Max.   :25.815  S07    : 8
                (Other):72

      anxiety
Min.   :21.0
1st Qu.:23.0
Median :25.0
Mean   :28.1
3rd Qu.:30.0
Max.   :49.0
```

Dans un premier temps, on cherche à savoir si le score moyen d'activité cérébrale (variable nommée score) diffère selon le temps d'exposition de l'image (visibility : 16ms ou 166ms).

### Question 1

*Dans le dispositif expérimental présenté ci-dessus, combien chaque sujet voit-il d'images avec un temps d'exposition court (16ms) ?*

**Réponse**

## Question 2

Proposez un modèle adapté pour répondre à cette problématique (vous donnerez l'expression mathématique de ce modèle dans le contexte de l'exercice et son nombre de paramètres).

## Réponse

Les commandes R suivantes présentent l'ajustement de deux modèles intégrant l'effet de la durée d'exposition à l'image sur le score d'activité cérébrale :

```
mod1 <- lm(score~visibility,data=dta)
mod2 <- lm(score~visibility+id,data=dta)
```

## Question 3

Des deux modèles ajustés par les commandes ci-dessus, lequel correspond à celui attendu en réponse à la question 2 ?

## Réponse

Les tables d'analyse de la variance des deux modèles sont données ci-après :

Analysis of Variance Table

```
Response: score
      Df Sum Sq Mean Sq F value Pr(>F)
visibility  1   62.9   62.87    2.496  0.117
Residuals 118 2971.6   25.18
```

Analysis of Variance Table

```
Response: score
      Df Sum Sq Mean Sq F value Pr(>F)
visibility  1   62.9   62.87    17.95 4.92e-05
id         14 2607.3  186.24    53.17 < 2e-16
Residuals
```

Dans la table d'analyse de la variance du 2ème modèle (mod2), les valeurs de la dernière ligne sont toutes manquantes (Df, Sum Sq et Mean Sq).

#### Question 4

En vous appuyant sur les deux tables d'analyse de la variance, donnez la valeur manquante dans la colonne *Sum Sq* de la table d'analyse de la variance du 2ème modèle (ne vous contentez pas de donner la valeur ; expliquez votre opération en quelques mots).

#### Réponse

Le tableau suivant donne la valeur estimée des paramètres du modèle mod2 :

```
summary(mod2)$coefficients
```

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	10.167874	0.661695	15.366414	1.62277e-28
visibility1	-0.723793	0.170849	-4.236456	4.92057e-05
idS02	1.496926	0.935778	1.599660	1.12708e-01
idS03	-2.012530	0.935778	-2.150650	3.38203e-02
idS04	-0.874332	0.935778	-0.934338	3.52294e-01
idS05	-0.553517	0.935778	-0.591505	5.55466e-01
idS07	-1.021838	0.935778	-1.091967	2.77371e-01
idS09	-6.781416	0.935778	-7.246825	7.76872e-11
idS10	4.500140	0.935778	4.808985	5.14178e-06
idS12	8.821074	0.935778	9.426465	1.28625e-15
idS13	1.377994	0.935778	1.472566	1.43888e-01
idS15	-4.659231	0.935778	-4.978995	2.54800e-06
idS16	3.385924	0.935778	3.618300	4.59630e-04
idS17	-3.940375	0.935778	-4.210803	5.42250e-05
idS19	-8.844952	0.935778	-9.451981	1.12819e-15
idS21	-8.009965	0.935778	-8.559689	1.08523e-13

#### Question 5

D'après le tableau précédent, quelle est la valeur estimée de la différence entre les scores moyens d'activité cérébrale pour un temps court d'exposition à une image et pour un temps long ?

#### Réponse

On soupçonne que l'activité cérébrale mesurée par EEG soit affectée par le niveau d'anxiété du sujet (mesuré sur une échelle numérique par la variable *anxiety*) au moment de l'expérience, et ce différemment selon le genre du sujet (variable *sex*).

### Question 6

Dans la problématique ci-dessus, quelle est la variable réponse et quelle est ou quelles sont les variables explicatives ? Donnez la nature quantitative ou catégorielle de ces variables.

### Réponse

### Question 7

Proposez un modèle adapté pour répondre à cette problématique (vous donnerez l'expression mathématique de ce modèle dans le contexte de l'exercice et son nombre de paramètres).

### Réponse

Le tableau ci-dessous donne les valeurs estimées des paramètres du modèle de la question 7 (dans ce tableau, le genre masculin est codé 1):

```
mod <- lm(score~sex*anxiety,data=dta)
summary(mod)$coefficients
```

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	4.219485	1.4711304	2.86819	4.90556e-03
sex1	8.647882	1.4711304	5.87839	4.06380e-08
anxiety	0.164939	0.0504306	3.27062	1.41326e-03
sex1:anxiety	-0.348540	0.0504306	-6.91128	2.76327e-10

### Question 8

*D'après le tableau ci-dessus, quelle est la valeur estimée du coefficient directeur de la droite reliant le score d'activité cérébrale et le score d'anxiété pour un sujet de sexe masculin ? En pratique, en concluez-vous que l'anxiété inhibe ou au contraire active l'activité cérébrale des sujets masculins ?*

### Réponse

On donne ci-après la table d'analyse de la variance du modèle reliant le score d'activité cérébrale au score d'anxiété (tous genres confondus) :

```
mod <- lm(score~anxiety,data=dta)
anova(mod)
```

Analysis of Variance Table

Response: score

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
anxiety	1	48.4	48.41	1.913	0.169
Residuals	118	2986.1	25.31		

### Question 9

*D'après la table d'analyse de la variance ci-dessus et le tableau donné avant la question 8, peut-on considérer que le score d'activité cérébrale est affecté de manière significative par l'anxiété ?*

### Réponse