

## TESTS STATISTIQUES

### Choix d'un test statistique & interprétation des résultats

*Mihaela Iancu*

#### Objectives

- ✓ Savoir les conditions d'applicabilité du test t de Student
- ✓ Enoncer les hypothèses du test t de Student pour échantillons indépendants
- ✓ Enoncer les hypothèses du test t de Student pour échantillons dépendants / paires
- ✓ Interprétation des résultats d'un test t de Student

#### Exercices

##### Scénario 1 :

On suppose qu'on a fait une étude pour évaluer l'association entre l'érosion dentaire et la consommation de boissons rafraîchissantes sans alcool, reflux gastrique, et obésité chez les jeunes adultes. 101 jeunes adultes qui viennent à une clinique dentaire à Cluj- Napoca (période de temps : mai 2017- janvier 2020) ont été choisis aléatoirement et inclus dans l'étude. La présence des lésions érosives a été déterminée par le BEWE (*engl. Basic Erosive Wear Examination*), un examen visuel au cours duquel toutes les dents, sauf les dents de sagesse, ont été examinées pour identifier des lésions dues aux acides. Les lésions érosives de la dent ont été classées sur une échelle ordinaire (de 0 à 3) représentant les degrés de sévérité (Tableau 1).

Tableau A. Description du score de BEWE

BEWE score	Description de l'aspect clinique du dent
0	Pas des lésions érosives
1	Perte initiale de texture de surface dentaire
2	Distinct défaut, la perte de tissu dur <50% de la surface dentaire
3	perte de tissus durs $\geq 50\%$ de la surface dentaire



Source : [https://fr.qwe.wiki/wiki/Tooth\\_wear](https://fr.qwe.wiki/wiki/Tooth_wear)

Le calcul du score du BEWE par patient se fait par addition des valeurs les plus élevées de chaque sextant (la cavité buccale est divisée en 6 groupes de dents appelées sextants).

Un patient ayant le score de BEWE  $\geq 9$  est considéré comme un patient souffrant d'érosion dentaire. Les résultats obtenus sont présentés dans le tableau ci-dessous :

**Tableau B. La distribution de l'Age en fonction de l'érosion dentaire chez les jeunes adultes**

Variable	Érosion dentaire <sup>(a)</sup> (n=30)	Absence d'érosion dentaire <sup>(b)</sup> (n=71)	p-valeur
Age (ans)	29,52±2,22	26,00±2,00	0,0281

Note. <sup>(a)</sup> BEWE $\geq 9$ , <sup>(b)</sup> BEWE $< 9$  ; Les variables quantitatives ont été décrites par la moyenne  $\pm$  écart-type et les variables qualitatives sont représentés par leur fréquences (absolues et relatives).

## Demandes

1. Créez un nouveau dossier **TP08NP** ou **N = votre nom, P = votre prénom** sur le bureau Bureau (Desktop) de Windows.
  1. Enregistrez ce document Word avec le nom **TP08MDFR** dans le dossier **TP08NP**.

**QUESTION DE RECHERCHE (Q1) :** En regardant les résultats donnés dans le **TABLEAU B**, au risque de 5%, on peut affirmer qu'il y a une **différence statistiquement significative** entre **les moyennes de l'AGE** chez les jeunes adultes **ayant érosions** dentaires et **ceux qui n'ont pas** ?? On suppose que les variances de l'AGE sont égales.

Pour répondre à la question de recherche (Q1), remplissez/répondez aux questions suivantes :

### QUE VOULONS-NOUS COMPARER?

2. Les jeunes adultes inclus dans cette étude ont été divisés par rapport à : un type du traitement/un facteur d'exposition (ex. genre)/ une maladie dentaire. Choisissez la bonne réponse et écrivez-le ci-dessous :
3. **L'AGE était une variable** : quantitative /qualitative dichotomique/ qualitative nominale/ qualitative ordinale. Choisissez la bonne réponse et écrivez-le ci-dessous :
4. **Les paramètres** testés pour identifier les différences en ce qui concerne l'AGE sont : les **moyennes/médians/variances/fréquences**. Choisissez la bonne réponse et écrivez-le ci-dessous :
6. Les groupes utilisés pour comparer l'AGE sont : des groupes **indépendants** ou **dépendants** ??
7. Combien des groupes sont utilisés pour comparer l'AGE : **une seule** groupe/**deux** groupes/**trois** groupes/**plusieurs** groupes.

### CHOISIR LE BON TEST STATISTIQUE :

8. Choisissez le test statistique approprié pour tester si la **moyenne de l'Age** chez les jeunes adultes souffrant d'**érosion dentaire** **diffère significativement** de la **moyenne de l'Age** chez les jeunes adultes qui **n'ont pas d'érosion dentaire**.
  - A. Test de Student-t pour groupes indépendantes avec des variances égales
  - B. Test de Student-t pour groupes indépendantes avec des variances inégales
  - C. Test de Student-t pour groupes dépendantes

**D. Test de Fisher****Copiez le nom du test et écrivez-le dans le carré ci-dessous :****FORMULATION DE L'HYPOTHESE NULLE DU TEST STATISTIQUE :**

9. Laquelle des hypothèses suivantes est l'hypothèse nulle associée à la question de recherche?
- A. Il n'y a pas de **différence statistiquement significative** entre **les moyennes de l'AGE des deux sous populations des jeunes adultes** (jeunes adultes ayant l'érosion dentaire et ceux qui n'ont pas).
  - B. Il n'y a pas de **différence statistiquement significative** entre **les moyennes de l'AGE des deux échantillons des jeunes adultes** (groupe des jeunes adultes ayant l'érosion dentaire et ceux qui n'ont pas).
  - C. Il n'y a pas de **différence statistiquement significative** entre **les médianes de l'AGE des deux sous populations des jeunes adultes** (jeunes adultes ayant l'érosion dentaire et ceux qui n'ont pas).
  - D. Il n'y a pas de **différence statistiquement significative** entre **les fréquences de l'AGE** chez les jeunes adultes ayant l'érosion dentaire et les jeunes adultes sans érosion dentaires (on parle ici des deux sous populations)

**Copiez la bonne réponse :****INTERPRETATION DES RESULTATS D'UN TEST STATISTIQUE a l'aide de la p-valeur**

10. Si le niveau de signification est  $\alpha = 0,05$ , copiez la bonne interprétation de la p-valeur du test qui a fait la comparaison entre les moyennes de l'AGE sur les deux populations (**voir le Tableau 2**):
- A.  **$p < 0,05$ , nous rejetons l'hypothèse nulle ( $H_0$ ) et nous sommes en faveur de l'hypothèse alternative ( $H_1$ )** : Il y a une différence statistiquement significative entre les **moyennes de l'AGE des deux sous populations des jeunes adultes** (jeunes adultes ayant l'érosion dentaire et ceux qui n'ont pas).
  - B.  **$p \geq 0,05$ , l'hypothèse nulle ( $H_0$ ) ne peut être rejetée**: Dans cette étude, on n'a pas trouvé une différence statistiquement significative entre les **moyennes de l'AGE des deux sous populations des jeunes adultes** (jeunes adultes ayant l'érosion dentaire et ceux qui n'ont pas). Formulation équivalente: les données ne fournissent pas des preuves suffisantes pour conclure qu'il y a une différence statistiquement significative entre les moyennes de l'AGE **des deux sous populations des jeunes adultes** (jeunes adultes ayant l'érosion dentaire et ceux qui n'ont pas).

**Copiez la bonne interprétation de la p-valeur :**

**APPLICATION DU TEST T DE STUDENT POUR DES ECHANTILLONS INDEPENDANTES**

**QUESTION DE RECHERCHE (Q2):** Au risque de 5%, on peut affirmer qu'il y a une **différence statistiquement significative** entre les **moyennes de l'IMC** chez les jeunes adultes ayant érosions dentaires et ceux qui n'ont pas ??

*Pour répondre à la question Q2, suives les étapes suivantes :*

1. Enregistrez le fichier **BDTP08\_MDFR** dans le dossier **TP08NP**.
2. Dans la feuille **Donnes\_Q2**, trier les données par rapport aux **EROSIONS DENTAIRES** (**voir les conseils**, page 2 ).
3. **Calculer les STATISTIQUES DESCRIPTIVES en utilisant l'option DATA ANALYSIS** (fr. Analyse de données) pour la variable IMC dans le **groupe de jeunes adultes ayant érosions dentaires**. (**voir Conseils du TP05**)
4. **Calculer les STATISTIQUES DESCRIPTIVES en utilisant l'option DATA ANALYSIS** (fr. Analyse de données) pour la variable IMC dans le **groupe de jeunes adultes qui n'ont pas d'érosions dentaires**. (**voir Conseils du TP05**)
5. **Décider (faire l'interprétation des statistiques descriptives) si la variable IMC suit la loi gaussienne** (normale) **dans chaque des deux groupes** de patients. (**voir Conseils du TP07**)
6. Dans la feuille **Donnes\_Q2 a cote de la cellule** Tester l'égalité des **variances** par le test de FISHER (**voir les conseils de ce TP, page 4** )
7. Tester l'égalité des **moyennes** par le test t de STUDENT pour **2 échantillons independantes** (**voir les conseils de ce TP, page 4**)
8. Remplir le **TABLEAU 1** trouve sur la feuille **Donnes\_Q2**.

**APPLICATION DU TEST T DE STUDENT POUR DES ECHANTILLONS DEPENDANTES (APPARIES)**

**QUESTION DE RECHERCHE (Q3):** Au risque de 5%, on peut affirmer qu'il y a une **différence statistiquement significative** entre les scores moyens de l'Indice BEWE chez les jeunes adultes **avant et après** le traitement ??

On suppose qu'on a fait une étude pour évaluer l'effet d'un certain traitement chez les s jeunes adultes ayant érosion dentaire. Pendant de 2 mois, chaque patient a reçu un traitement dentaire contenant du fluor et calcium. L'évaluation de l'effet du traitement a été réalisé en mesurant le score BEWE avant et après 2 mois du traitement.

*Pour répondre à la question Q3, suives les étapes suivantes :*

1. Allez sur la page **Donnes\_Q3**.
2. **Calculer les STATISTIQUES DESCRIPTIVES en utilisant l'option DATA ANALYSIS** pour la variable **Indice BEWE avant** du traitement.

3. Calculer les **STATISTIQUES DESCRIPTIVES** en utilisant l'option **DATA ANALYSIS** pour la variable **Indice BEWE après** le traitement.
4. **Décider (faire l'interprétation des statistiques descriptives)** si les deux variables (**Indice BEWE avant & Indice BEWE après**) suivent la lois gaussienne (normale).
5. Tester l'égalité **des moyennes** par le test **t de Student** pour **2 échantillons dépendantes/appariés** (**voir les conseils de ce TP, page 5**).
6. Remplir le **TABLEAU 2** trouve sur la feuille **Donnes\_Q3**.

### Probleme récapitulatif/révision

*On considère la même base de données du Scenario 1.*

1. Déterminer l'**intervalle de confiance a 95%** pour la **moyenne** de l'**IMC** chez les **patients souffrant d'érosions dentaires**.
2. Déterminer l'**intervalle de confiance a 95%** pour la **moyenne** de l'**IMC** chez les **patients qui ne souffre pas d'érosions dentaires**.
3. **En utilisant les résultats obtenus au points 1 & 2**, au risque de 5%, on peut affirmer qu'il y a une différence significative entre les moyennes de l'IMC dans les deux populations? (*Conseils : si les intervalles 95% IC sont disjointes  $\Leftrightarrow$  différence significative ; au cas contraire, on peut pas dire s'il y a une différence significative*). Ecrivez la réponse pour le point 3 sur votre page de calcul.
4. Faites un graphique pour la répartition des sujets par rapport aux **érosions dentaires**.
5. Calculer la **probabilité** pour qu'un individu pris au hasard ait **érosions dentaires**.

### Rappelez-vous que....

1. L'**HYPOTHESE NULLE (H0)** du test t de **STUDENT** affirme que : Il n'y a pas de différence statistiquement significative entre les **MOYENNES** de la variable d'intérêt dans les deux populations
2. L'**HYPOTHESE NULLE (H1)** du test t de **STUDENT** affirme que : Il y a une différence statistiquement significative entre les **MOYENNES** de la variable d'intérêt dans les deux populations
3. Le test de **STUDENT** sera utilisé pour tester l'égalité entre les **MOYENNES** en sachant que les distributions se rapproche de la loi Normale (ou de Student)
4. SI la p-valeur du test de **STUDENT**  $< 0.05 \Rightarrow$  on rejette **H0**  $\Rightarrow$  nous sommes en faveur du **H1**  $\Rightarrow$  au risque de 5%, il y a une différence statistiquement significative entre les **MOYENNES** de la variable d'intérêt dans les deux populations
5. Le test de **FISHER** sera utilisé pour tester l'égalité entre les **VARIANCES** pour deux échantillons **INDEPENDANTES**
6. SI la p-valeur du test de **FISHER**  $< 0.05 \Rightarrow$  on rejette **H0**  $\Rightarrow$  nous sommes en faveur du **H1**  $\Rightarrow$  au risque de 5%, il y a une différence statistiquement significative entre les **VARIANCES** de la variable d'intérêt dans les deux populations