



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara



Tests statistiques: concepts, décisions, erreurs statistiques.

Teste statistiques usuels(I): le test t-Student

Cours 6

Dr. Mirela FRANDES



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

Tests statistiques

	Pour comparer	Pour vérifier
De conformité	Une valeur obtenue avec une valeur théorique	
D'omogenite	Les valeurs de 2 paramètres de dispersion	que 2 échantillons proviennent de la même population/ou que les deux échantillons ne sont pas différents
De signification	Les valeurs de 2 moyennes ou proportions	que 2 échantillons proviennent de la même population/ou que les deux échantillons ne sont pas différents
De concordance	une distribution expérimentale avec une distribution théorique ou 2 distributions expérimentales	
D'indépendance	l'indépendance des séries de valeurs	
De corelation	La signification des paramètres	pour prouver l'existence d'une liaison entre 2 variables



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

Tests statistiques

-
- **Hypothèses statistiques** - postule l'existence d'une différence entre groupes ou d'une association entre paramètres

La signification statistique est un élément très important qui assure que le résultat obtenu a une forte chance d'être réel et non pas d'être obtenu par hasard

- Significative du point de vue statistique \neq significative du point de vue biomédicale
-



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

Différences

1. Insignifiantes

- A forte probabilité d'apparaître par hasard
- La cause est la variabilité de l'échantillonnage
- Les deux échantillons appartiennent à la même population

2. Significatives

- La probabilité d'apparaître par hasard est faible
 - Les différences sont attribuées à d'autres causes
-



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

Les étapes d'utilisation des tests statistiques

1. Définir les mesures (variables) qu'on compare (en fonction de leur type on va choisir le test statistique)
2. Établir l'hypothèse nulle (H_0) et l'hypothèse alternative H_a (ou H_1)
3. Établir un niveau de signification (α) (ce qui donne le niveau de confiance)
4. Choisir le test statistique approprié pour tester l'hypothèse nulle (H_0)
5. Calculer la valeur du test statistique à l'aide des données de l'échantillon; la probabilité p
6. Comparer p et α
7. Formuler la conclusion



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

Les étapes d'utilisation des tests statistiques

- p = probabilité que :
 - H_0 est vrai
 - Les différences ne sont pas significatives
 - On est dans la région d'acceptation de H_0
- 7. Formuler la décision (en fonction du p)
 - $p \geq \alpha$ - on accepte H_0
 - $p < \alpha$ - on rejette H_0



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

Les hypothèses statistiques

L'hypothèse nulle (H_0)

- Est une hypothèse de non différence : il n'y a pas de différences significatives entre échantillons
- les échantillons sont égaux **du point de vue statistique**

L'hypothèse alternative (H_a)

- Il y a des différences significatives entre les échantillons
- Elle s'oppose à l'hypothèse nulle
- les échantillons sont différent **du point de vue statistique**



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

Les hypothèses statistiques

L'hypothèse nulle (H_0)

- **Il n'y a pas** de différence entre 2 traitements, entre 2 groups
- **Il n'y a pas** un relation/lien/association entre 2 caractéristiques: e.g., Facteur de risque - maladie

L'hypothèse alternative (H_a)

- **Il y a** un différence (entre 2 traitements, entre 2 groups)
- **Il y a** un relation/association entre 2 caractéristiques: e.g., Facteur de risque - maladie



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

Tests statistiques

1. une valeur moyenne obtenue avec une valeur théorique
2. 2 valeurs moyennes pour 2 séries indépendants ou paires
3. >2 valeurs moyennes
4. 2 dispersions
5. >2 (plus du 2) dispersions
6. une distribution expérimentale avec une distribution théorique



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

Exemple

- Dans un essai thérapeutique (l'administration d'un médicament)
 - l'hypothèse nulle
 - correspond à l'absence d'effet du traitement étudié
 - l'hypothèse alternative
 - l'effet du traitement n'est pas nul – le médicament est efficace



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

Le seuil de signification (α)

- Est la valeur de la probabilité à partir de laquelle nous commençons à prendre en considération les différences comme significatives
- Il a été convenu que le risque acceptable d'erreur α est de 5%
- Ainsi, devant une différence observée, on conclura à l'existence d'une réelle différence seulement si le risque de se tromper pris en faisant cette conclusion est inférieur à 5%, c'est-à-dire, si la valeur de p donnée par le test est inférieure au seuil de 0.05
- **Le risque α est le risque/la probabilité de commettre une erreur statistique**



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

Les erreurs statistiques

- Type I
 - α
 - L'erreur qui consiste à rejeter une hypothèse vraie
 - L'hypothèse nulle (H_0) = est vrai, mais pas acceptée
- Type II
 - β
 - L'erreur qui consiste en acceptant une hypothèse fausse
 - L'hypothèse nulle (H_0) = est false, mais acceptée



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

Les erreurs statistiques

		Situation réelle	
		H0 vrai	H0 faux
La décision	Accepté H0	$p=1-\alpha$	$p=\beta$
	Rejeté H0	$p=\alpha$	$p=1-\beta$



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

Les caractéristiques du test

Niveau de confiance ($1 - \alpha$)

- capacité d'accepter une hypothèse quand elle est vrai

Puissance du test ($1 - \beta$)

- capacité de rejeter une hypothèse quand elle est faux



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

La décision statistique

p	Les différences	Les différences	L'hypothèse nulle(H_0)
$> 0,05$	négligeables	N	acceptée
$< 0,05$	significative	S	rejetée
$< 0,01$	très significative	TS	rejetée
$< 0,001$	extrêmement significative	ES	rejetée



Tests paramétriques et non-paramétriques

1. Tests paramétriques

- Pour les variable avec une **distribution connue** ou normale
- Sont plus puissants que les tests non paramétriques

2. Tests non paramétriques

- Pour une **distribution inconnue**
 - On l'applique quand les conditions d'applications des autres méthodes ne sont pas satisfaites.
 - Les probabilités obtenues par la plupart des tests non paramétriques sont des probabilités exactes
 - Pour des échantillons de taille très faible jusqu'à $n = 6$
-



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

Tests statistiques

	Pour comparer	Pour vérifier
De conformité	Une valeur obtenue avec une valeur théorique	
D'omogenite	Les valeurs de 2 paramètres de dispersion	que 2 échantillons proviennent de la même population/ou que les deux échantillons ne sont pas différents
De signification	Les valeurs de 2 moyennes ou proportions	que 2 échantillons proviennent de la même population/ou que les deux échantillons ne sont pas différents
De concordance	une distribution expérimentale avec une distribution théorique ou 2 distributions expérimentales	
D'indépendance	l'indépendance des séries de valeurs	
De corelation	La signification des paramètres	pour prouver l'existence d'une liaison entre 2 variables



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

Selon la constitution des échantillons

	Quoi?	Exemple
Pour échantillons indépendants	les observations sont indépendantes à l'intérieur des groupes et d'un groupe à l'autre	TA aux hommes et aux femmes
Pour échantillons appariés	la situation des mesures répétées sur les mêmes sujets	TA avant et après un médicament



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

Tests statistiques

2. Établir un niveau de signification (α)

α est une probabilité de rejet incorrect de H_0 quand elle est vraie

Valeurs traditionnelles:

- $\alpha = 0,05$ (5% erreur)
- $\alpha = 0,01$ (1% erreur)
- $\alpha = 0,001$ (0.1 % erreur)

3. Établir le test statistique approprié pour tester l'hypothèse nulle (H_0)

si l'échantillon est:

large ($n \geq 30$) on utilise la statistique Z (test normale)

petit ($n < 30$) on utilise la statistique t (test t-Student) avec $n-1$ degrés de liberté



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

Interpréter le résultat

- **La décision avec *p-value***
 - *p-value* est la probabilité d'obtenir, quand H_0 est supposée vraie, un résultat pour la statistique du test plus extrême que le résultat observe
 - Si $p\text{-value} < \alpha (=0,05) \Rightarrow$ on accepte H_a et on rejete H_0
 - Si $p\text{-value} > \alpha (=0,05) \Rightarrow$ on accepte H_0
- **$p = 0.42390353 > \alpha (=0,05) \Rightarrow$ on accepte H_0 : il n'existe pas une différence statistiquement significative entre la valeur du cholestérol avant et après le traitement**



Tests statistiques usuels (2 moyennes)

- 2 séries indépendants
 - $H_0 : m_1 = m_2$
 - condition: $s_1 = s_2$
 - Test t-Student pour échantillons indépendants
(test de signification, paramétrique)
- 2 séries paires
 - $H_0 : m_1 = m_2$
 - Test t-Student pour différence
(test de signification, paramétrique)



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

Test t-Student

(2 échantillons indépendants)

- Conditions d'application:
 - les variances des échantillons sont connues
 - C'est possible $n_1 \neq n_2$
- l'hypothèse nulle (H_0) - Il n'y a pas de différence entre les deux moyennes (la différence entre les deux moyennes est égale à zéro)
- l'hypothèse alternative (H_1) du test bilatéral - Il y a une différence entre les deux moyennes (la différence entre les deux moyennes n'est pas égale à zéro)



UMFT

Universitatea de
Medicină și Farmacie
„Victor Babeș”
din Timișoara

Test t-Student

(2 échantillons indépendants)

- **H0 : $\mu_1 = \mu_2$**
- **H1 : $\mu_1 \neq \mu_2$**

$$t = \frac{m_1 - m_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

- n_1, n_2 = les tailles des échantillons
- m_1, m_2 = les moyennes des échantillons
- s_1, s_2 = les variances

□ On cherche dans la tableau de valeurs de la distribution t, la valeur critique en fonction de alpha et de nombre de degrés de liberté

- Si la valeur t obtenue est plus grande de la valeur du tableau: H0 est rejeté
- Si la valeur t obtenue est plus petite de la valeur du tableau: H0 est accepté