Principaux tests statistiques pour échantillons de petites tailles

Etienne Dantan

(Etienne.Dantan@univ-nantes.fr)

Master 2 BBRT, DCPS et SANH

18 Octobre 2013

Plan

Plan

Contexte

- On considère deux populations $\mathcal{P}_A$ et $\mathcal{P}_B$ (ou plus) desquelles sont extraits deux échantillons de tailles n_A et n_B.
- A partir de ces observations, on cherche à savoir si les caractéristiques des deux populations peuvent être considérées comme égales, ou bien si elles semblent être différentes.

Les acquis (1)

- 2 types de variables aléatoires

1- *v.a. quantitatives :*

→ **v.a. discontinues ou discrètes**
x_i : *nombre fini* (ex. : lancé de dé) ou *nombre infini dénombrable* (ex. : nombre de petits par portée pour une espèce animale donnée, nombre de mutations dans une séquence ADN, etc.)

→ **v.a. continues**
x_i : *nombre infini non dénombrable* (ex. : taille ou âge d'un individu tiré au hasard)

Les acquis (2)

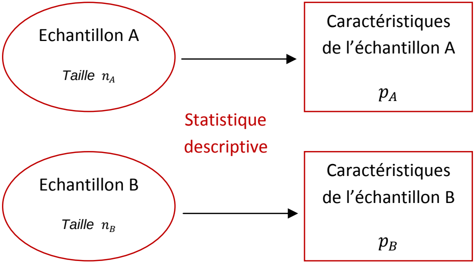
2- *v.a. qualitatives :*

$\rightarrow$ **v.a. ordinales**
x_i : *valeurs non numériques pouvant être ordonnées* (ex. : intensité d'une douleur (faible=1, moyenne=2 ou forte=3))

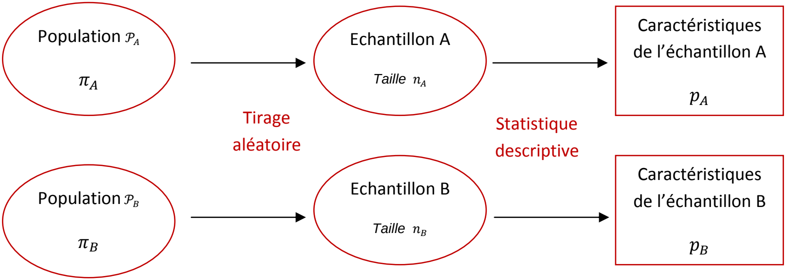
$\rightarrow$ **v.a. catégorielles ou nominales**
x_i : *valeurs non numériques ne pouvant pas être ordonnées* (ex. : malade=1 ou non malade=0, fumeurs=1 ou non fumeurs=0)

$\Rightarrow$ en fonction de la question d'intérêt : étude quantitative ou fréquentielle

Exemple - Comparaison de 2 fréquences



Exemple - Comparaison de 2 fréquences



$$H_0 : \pi_A = \pi_B = \pi$$

Exemple - Comparaison de 2 fréquences

Conclusion

- rejet de H_0 : la différence entre π_A et π_B n'est pas due au hasard,

  i.e. échantillons A et B semblent représentatifs de 2 populations distinctes

- non rejet de H_0 : la différence entre π_A et π_B semble due aux fluctuations d'échantillonnage,

  i.e. nous n'avons pas pu mettre en évidence de différence significative

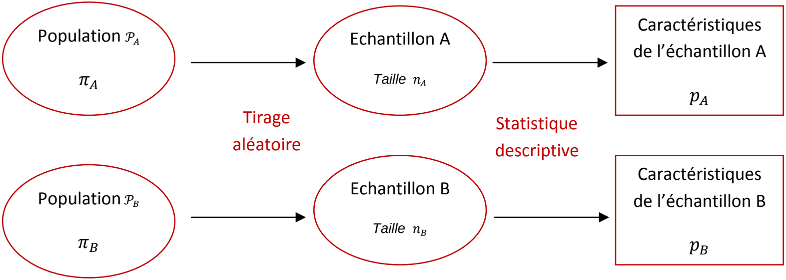
$\Rightarrow$ attention : il existe **toujours** une probabilité de se tromper

α : probabilité de se tromper en rejetant H_0

β : probabilité de se tromper en ne rejetant pas H_0, i.e. en rejetant H_1

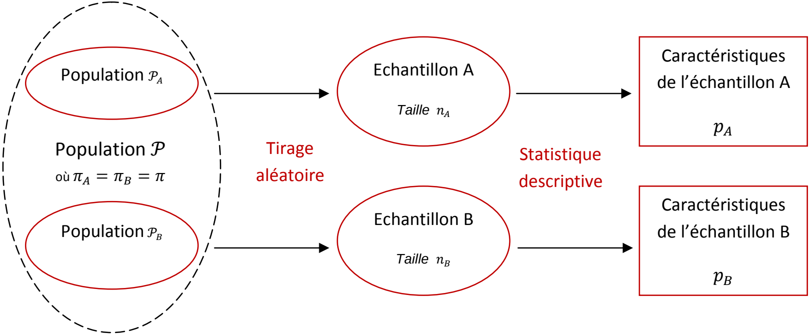
Exemple - Comparaison de 2 fréquences

Rejet de H_0 au risque de 1ère espèce α



Exemple - Comparaison de 2 fréquences

Non rejet de H_0 au risque de 1ère espèce α



Les $\neq$ cas de figure (1)

- 1 Variable quantitative continue à expliquer

  en fonction d'une autre variable qualitative explicative à deux modalités

  $\Rightarrow$ comparaison de deux moyennes

$\rightarrow$ Test de Student de comparaison de deux moyennes issues de 2 populations **indépendantes** représentées par 2 **grands** échantillons (test paramétrique : approximation normale)

$\rightarrow$ Test de Mann-Whitney de comparaison de 2 distributions de 2 v.a. issues de 2 populations **indépendantes** représentées par 2 **petits** échantillons (test non-paramétrique)

Les $\neq$ cas de figure (2)

- 1 Variable quantitative continue à expliquer

  en fonction d'une autre variable qualitative explicative à + de 2 modalités

  $\Rightarrow$ comparaison de + de 2 moyennes

$\rightarrow$ ANOVA
à partir de + de 2 populations **indépendantes** représentées par des **grands** échantillons

$\rightarrow$ Test de Kruskal-Wallis de comparaison de + de 2 distributions de v.a. issues de + de 2 populations **indépendantes** représentées par des **petits** échantillons (test non-paramétrique)

Les $\neq$ cas de figure (3)

- 1 Variable quantitative continue à expliquer

  en fonction d'une autre variable qualitative explicative à deux modalités

  $\Rightarrow$ comparaison de deux moyennes

$\rightarrow$ Test de Student de comparaison d'une différence à 0 pour 2 populations **appariées** représentées par 2 **grands** échantillons (test paramétrique)

$\rightarrow$ Test de Wilcoxon de comparaison de 2 distributions de 2 v.a. issues de 2 populations **appariées** représentées par 2 **petits** échantillons (test non-paramétrique)

Les $\neq$ cas de figure (4)

- 1 Variable qualitative à expliquer

  en fonction d'une autre variable qualitative explicative à deux modalités

  $\Rightarrow$ comparaison de deux pourcentages

$\rightarrow$ Test de comparaison de deux fréquences
issues de 2 populations **indépendantes** représentées par 2 **grands** échantillons (test paramétrique : approximation normale)

$\rightarrow$ Test exact de Fisher de comparaison de deux fréquences
issues de 2 populations **indépendantes** représentées par 2 **petits** échantillons (test non-paramétrique)

Les $\neq$ cas de figure (5)

- 1 Variable qualitative à expliquer

  en fonction d'une autre variable qualitative explicative à deux modalités

  $\Rightarrow$ comparaison de deux pourcentages

$\rightarrow$ Test de Mc Nemar de comparaison de deux fréquences
issues de 2 populations **appariées** représentées par 2 **grands** échantillons (test paramétrique)

$\rightarrow$ Test de Mc Nemar de comparaison de deux fréquences
issues de 2 populations **appariées** représentées par 2 **petits** échantillons (calcul exact proba critique basé sur la loi binomiale)

Les $\neq$ cas de figure (6)

- 1 Variable quantitative à expliquer

  en fonction d'une autre variable quantitative explicative

$\rightarrow$ Coefficient de corrélation

Construction d'un test

$\Rightarrow$ 6 étapes

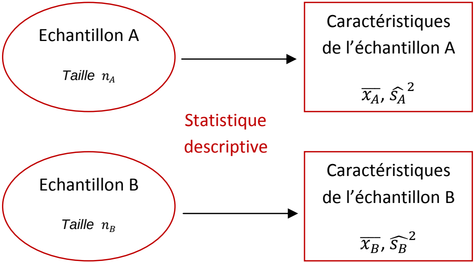
1. Définition des hypothèses nulle H_0 et alternative H_1
2. Choix du test, écriture de la statistique de test et de sa loi sous H_0
3. Choix du niveau de signification α, Définition de la région critique (RC)
4. Calcul de la statistique de test
5. Conclusion statistique : rejet ou non de H_0
6. Conclusion épidémiologique

Plan

Plan

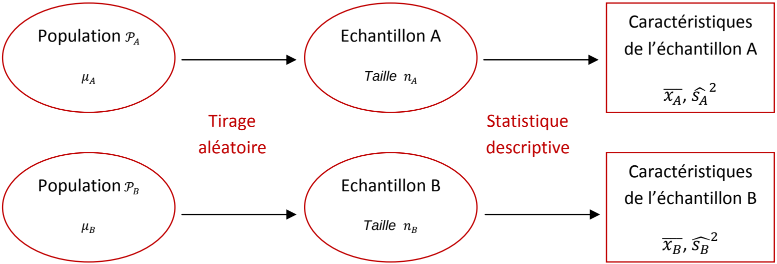
Principe

On observe 2 échantillons A et B indépendants (i.e. individus $\neq$ dans les 2 groupes)



Comparaison de 2 moyennes

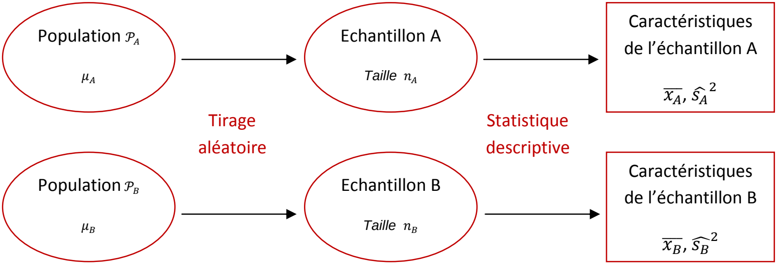
Raisonnement par l'absurde :



$$H_0 : \mu_A = \mu_B \quad ?$$

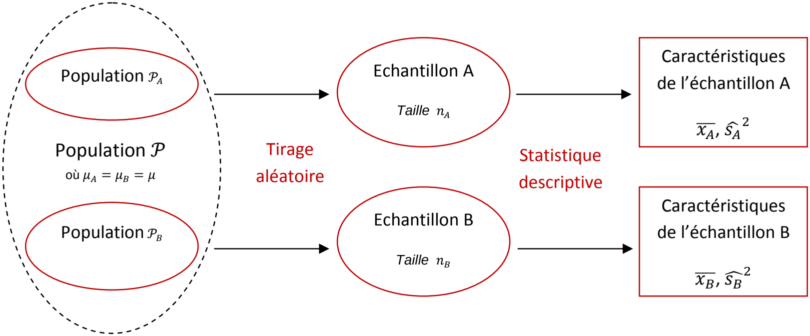
Comparaison de 2 moyennes

Si **rejet** de H_0, alors les 2 échantillons semblent issus de 2 populations distinctes :



Comparaison de 2 moyennes

Si **non rejet** de H_0, alors les 2 échantillons semblent issus d'une seule et même population :



Contexte

- Définition des variables aléatoires :
  - X_A : v.a. continue dans la population $\mathcal{P}_A$ de moyenne μ_A.

    On observe un échantillon de taille n_A $\{x_{A,1}, x_{A,2}, \ldots, x_{A,n_A}\}$.
  - X_B : v.a. continue dans la population $\mathcal{P}_B$ de moyenne μ_B.

    On observe un échantillon de taille n_B $\{x_{B,1}, x_{B,2}, \ldots, x_{B,n_B}\}$.

$\Rightarrow$ Test usuel :

Comparaison de 2 moyennes

Contexte

- Définition des variables aléatoires :
  - X_A : v.a. continue dans la population $\mathcal{P}_A$ de moyenne μ_A.
    On observe un échantillon de taille n_A $\{x_{A,1}, x_{A,2}, ..., x_{A,n_A}\}$.
  - X_B : v.a. continue dans la population $\mathcal{P}_B$ de moyenne μ_B.
    On observe un échantillon de taille n_B $\{x_{B,1}, x_{B,2}, ..., x_{B,n_B}\}$.

$\Rightarrow$ Test usuel : **Test paramétrique de STUDENT**

Limites

- Test nécessitant des effectifs de grandes tailles (n$\geq$30)
  $\rightarrow$ car théorème central limite valide

- Test de Student avec des petits échantillons existe, **MAIS hypothèses à vérifier** :
  - les 2 variables distribuées selon une loi normale
  - les variances des 2 variables égales

  $\Rightarrow$ en générale Manque de puissance pour montrer les 2 propositions précédentes

$\Rightarrow$ Solution :

Limites

- Test nécessitant des effectifs de grandes tailles (n$\geq$30)
  $\rightarrow$ car théorème central limite valide

- Test de Student avec des petits échantillons existe, **MAIS hypothèses à vérifier** :
  - les 2 variables distribuées selon une loi normale
  - les variances des 2 variables égales

  $\Rightarrow$ en générale Manque de puissance pour montrer les 2 propositions précédentes

$\Rightarrow$ Solution : TEST NON PARAMETRIQUE

Principes des tests non paramétriques

- Aucune hypothèse sur la distribution des v.a.
- Tests souvent basés sur la notion de rangs
  $\longrightarrow$ si les distributions entre groupes sont $\neq$, les rangs sont $\neq$
- Exemple :
  - Groupe A (n=3) : 1, 5, 3
  - Groupe B (n=3) : 7, 6, 10
- Rangs :
  - Groupe A (n=3) : 1, 3, 2
  - Groupe B (n=3) : 5, 4, 6
- Sommes des rangs :
  - Groupe A (n=3) : 6
  - Groupe B (n=3) : 15

Principes des tests non paramétriques

- Tests paramétriques :
  $\longrightarrow$ exigent une spécification de la forme de la distribution
- Tests non paramétriques :
  $\longrightarrow$ pas de référence à une répartition particulière
  $\longrightarrow$ peuvent s'appliquer à des petits échantillons
- Avantages/Inconvénients :
  $\longrightarrow$ tests non paramétriques théoriquement moins puissants que tests paramétriques
  $\longrightarrow$ études ont montré : exactitude des tests non paramétriques sur des grands échantillons n'est que légèrement inférieure à celle des tests paramétriques
  $\longrightarrow$ tests non paramétriques sont bcp plus exacts sur des petits échantillons

Comparaison de 2 moyennes :

Test non paramétrique de Mann-Whitney

- permet de comparer la distribution de 2 v.a. observées à partir de 2 échantillons indépendants (A et B)

- Définition des v.a. :
  - X_A : v.a. continue dans le groupe A de taille n_A
  - X_B : v.a. continue dans le groupe B de taille n_B

- Par convention, on définit le groupe comme le plus petit échantillon ($n_A \leq n_B$)

Comparaison de 2 moyennes :

Test non paramétrique de Mann-Whitney

- Choix des hypothèses
  - H_0 : X_A et X_B ont la même distribution
  - H_1 : X_A et X_B n'ont pas la même distribution
- Somme des rangs
  - R_A : somme des rangs occupés par les valeurs observées de X_A
  - R_B : somme des rangs occupés par les valeurs observées de X_B
- Si des ex-aequos entre des valeurs, on affecte la moyenne des rangs
  (si les ex-aequos sont du même groupe, rien ne change)

Comparaison de 2 moyennes :

Test non paramétrique de Mann-Whitney

- Statistique de test :

$$M = min(M_A, M_B)$$

avec

$$M_A = n_A n_B + n_A(n_A + 1)/2 - R_A$$

et

$$M_B = n_A n_B - U_A = n_A n_B + n_B(n_B + 1)/2 - R_B$$

- Choix du seuil α (5%) et définition de la région critique : 2 cas de figure
  - $n_A \leq 10$
  - $n_A > 10$

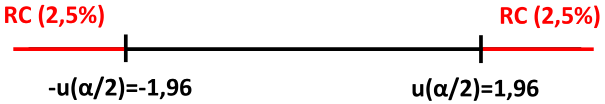
Comparaison de 2 moyennes :
Test non paramétrique de Mann-Whitney

- Si $n > 10$, alors sous H_0 :

$$M \sim \mathcal{N}(n_A n_B/2, \sqrt{n_A n_B(n_A + n_B + 1)/12})$$

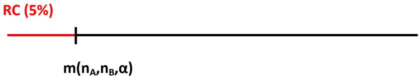
$$U = \frac{M - n_A n_B/2}{\sqrt{n_A n_B(n_A + n_B + 1)/12}} \sim \mathcal{N}(0,1)$$

$\rightarrow$ test bilatéral, loi normale



Comparaison de 2 moyennes :
Test non paramétrique de Mann-Whitney

- Si $n \leq 10$, alors :



- Lecture dans la table de la valeur critique $m(n_A; n_B, \alpha)$

Comparaison de 2 moyennes :

Test non paramétrique de Mann-Whitney

Critical Values of the Mann-Whitney U
(Two-Tailed Testing)

n_2	α	n_1																	
		3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3	.05	--	0	0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8
	.01	--	0	0	0	0	0	0	0	0	1	1	1	2	2	2	2	3	3
4	.05	--	0	1	2	3	4	4	5	6	7	8	9	10	11	11	12	13	14
	.01	--	--	0	0	0	1	1	2	2	3	3	4	5	5	6	6	7	8
5	.05	0	1	2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20
	.01	--	--	0	1	1	2	3	4	5	6	7	7	8	9	10	11	12	13
6	.05	1	2	3	5	6	8	10	11	13	14	16	17	19	21	22	24	25	27
	.01	--	0	1	2	3	4	5	6	7	9	10	11	12	13	15	16	17	18
7	.05	1	3	5	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34
	.01	--	0	1	3	4	6	7	9	10	12	13	15	16	18	19	21	22	24
8	.05	2	4	6	8	10	13	15	17	19	22	24	26	29	31	34	36	38	41
	.01	--	1	2	4	6	7	9	11	13	15	17	18	20	22	24	26	28	30
9	.05	2	4	7	10	12	15	17	20	23	26	28	31	34	37	39	42	45	48
	.01	0	1	3	5	7	9	11	13	16	18	20	22	24	27	29	31	33	36
10	.05	3	5	8	11	14	17	20	23	26	29	33	36	39	42	45	48	52	55
	.01	0	2	4	6	9	11	13	16	18	21	24	26	29	31	34	37	39	42
11	.05	3	6	9	13	16	19	23	26	30	33	37	40	44	47	51	55	58	62
	.01	0	2	5	7	10	13	16	18	21	24	27	30	33	36	39	42	45	48
12	.05	4	7	11	14	18	22	26	29	33	37	41	45	49	53	57	61	65	69
	.01	1	3	6	9	12	15	18	21	24	27	31	34	37	41	44	47	51	54
13	.05	4	8	12	16	20	24	28	33	37	41	45	50	54	59	63	67	72	76
	.01	1	3	7	10	13	17	20	24	27	31	34	38	42	45	49	53	56	60
14	.05	5	9	13	17	22	26	31	36	40	45	50	55	59	64	67	74	78	83
	.01	1	4	7	11	15	18	22	26	30	34	38	42	46	50	54	58	63	67
15	.05	5	10	14	19	24	29	34	39	44	49	54	59	64	70	75	80	85	90
	.01	2	5	8	12	16	20	24	29	33	37	42	46	51	55	60	64	69	73
16	.05	6	11	15	21	26	31	37	42	47	53	59	64	70	75	81	86	92	98
	.01	2	5	9	13	18	22	27	31	36	41	45	50	55	60	65	70	74	79
17	.05	6	11	17	22	28	34	39	45	51	57	63	67	75	81	87	93	99	105
	.01	2	6	10	15	19	24	29	34	39	44	49	54	60	65	70	75	81	86
18	.05	7	12	18	24	30	36	42	48	55	61	67	74	80	86	93	99	106	112
	.01	2	6	11	16	21	26	31	37	42	47	53	58	64	70	75	81	87	92
19	.05	7	13	19	25	32	38	45	52	58	65	72	78	85	92	99	106	113	119
	.01	3	7	12	17	22	28	33	39	45	51	56	63	69	74	81	87	93	99
20	.05	8	14	20	27	34	41	48	55	62	69	76	83	90	98	105	112	119	127
	.01	3	8	13	18	24	30	36	42	48	54	60	67	73	79	86	92	99	105

Comparaison de 2 moyennes :

Test non paramétrique de Mann-Whitney

- Si $u \in RC \rightarrow p < \alpha$:
  - Rejet de H_0 car moins de 5% de chances qu'elle soit vraie
  - Il semble que les 2 distributions soient différentes
- Si $u \notin RC \rightarrow p > \alpha$:
  - Non rejet de H_0 car plus de 5% de chances qu'elle soit vraie
  - On ne peut pas montrer que les 2 distributions soient différentes

Comparaison de 2 moyennes :

Test non paramétrique de Mann-Whitney

- Remarques de vocabulaire :
  - Ce test est aussi appelé Mann-Withney/Wilcoxon
  - On peut aussi voir “test de Wilcoxon pour échantillons indépendants”
  - A éviter mais rencontré dans la littérature : “non-parametric t-test”

- Problèmes des faibles effectifs :
  - Du point de vue statistique :
    - $\rightarrow$ Test possible à partir de 3 sujets par groupe
  - Du point de vue méthodologique :
    - $\rightarrow$ Résultats très peu robustes : un sujet supplémentaire peut tout changer
    - $\rightarrow$ Résultats très peu puissants : attention à l'interprétation du non-rejet de H_0

Comparaison de 2 moyennes :

Exemple

- Des biologistes s'intéressent à l'effet d'un nouveau médicament en phase pré-clinique. Ils testent sur 5 souris le médicament de référence (médicament A) et sur 8 autres souris la nouvelle molécule (médicament B).
- 13 prélèvements sanguins sont réalisés. Les concentrations (en micro/L) obtenues sont présentées dans le tableau suivant :

Concentration								
Médicament A	25	12	7	9	18			
Médicament B	17	19	14	6	16	20	22	13

X_A : concentration du médicament A dans le sang
X_B : concentration du médicament B dans le sang

Comparaison de 2 moyennes :
Exemple

On cherche à répondre à la question suivante :
"Les concentrations de médicament B dans le sang sont-elles significativement différentes de celles du médicament A ?" (risque d'erreur de 1ère espèce fixé à 5%) ?

- Quel test statistique doit-on réaliser ?

Comparaison de 2 moyennes :
Exemple

On cherche à répondre à la question suivante :
"Les concentrations de médicament B dans le sang sont-elles significativement différentes de celles du médicament A ?" (risque d'erreur de 1ère espèce fixé à 5%) ?

- Quel test statistique doit-on réaliser ?

Test de Mann-Whitney pour échantillons indépendants

Conditions $\Rightarrow$ échantillons de petites tailles ($n < 30$)

Test d'une $\neq$ $\Rightarrow$ test bilatéral

Comparaison de 2 moyennes :
Exemple

- Quelle est l'hypothèse nulle H_0 de ce test ?
- Quelle est l'hypothèse alternative H_1 de ce test ?

Comparaison de 2 moyennes :
Exemple

- Quelle est l'hypothèse nulle H_0 de ce test ?
- Quelle est l'hypothèse alternative H_1 de ce test ?

H_0 : les distributions des concentrations dans le sang des médicaments A et B sont identiques, autrement dit X_A et X_B ont la même distribution

H_1 : les distributions issues des 2 concentrations des 2 médicaments ne sont pas identiques, c'est-à-dire que X_A et X_B n'ont pas la même distribution

Comparaison de 2 moyennes :
Exemple

Remarque : on exclue les valeurs nulles

Rang								
Médicament A	13	4	2	3	9			
Médicament B	8	10	6	1	7	11	12	5

- Somme des rangs $R_A = 31$
- Somme des rangs $R_B = 60$

$\Rightarrow$ Statistique de test

$$M = min(5*8+5*6/2-31, 3*3+3*4/2-12)$$
$$= min(24, 16) = 16$$

Comparaison de 2 moyennes :
Exemple

$n_A = 5 < 10 \Rightarrow$ on lit dans la table la valeur critique définissant la région critique.

Au seuil $\alpha = 0.05$, pour $n_A = 5$ et $n_B = 8$, on a pour valeur critique la valeur 6

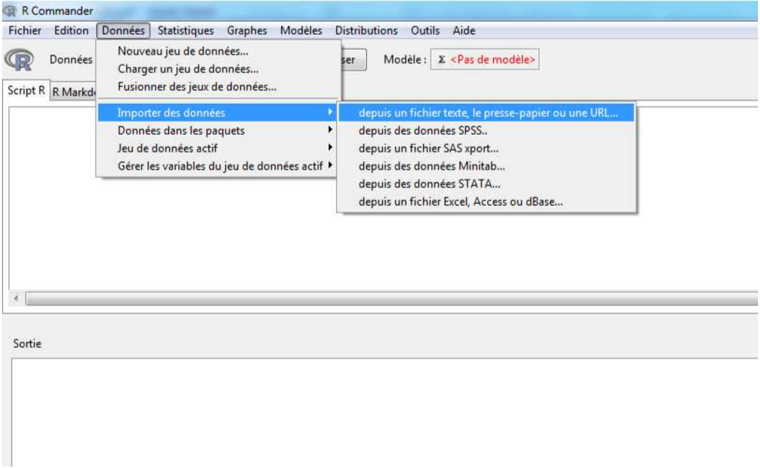
$M \notin$ à la région critique

$\Rightarrow$ Non Rejet H_0 au risque 5% de se tromper

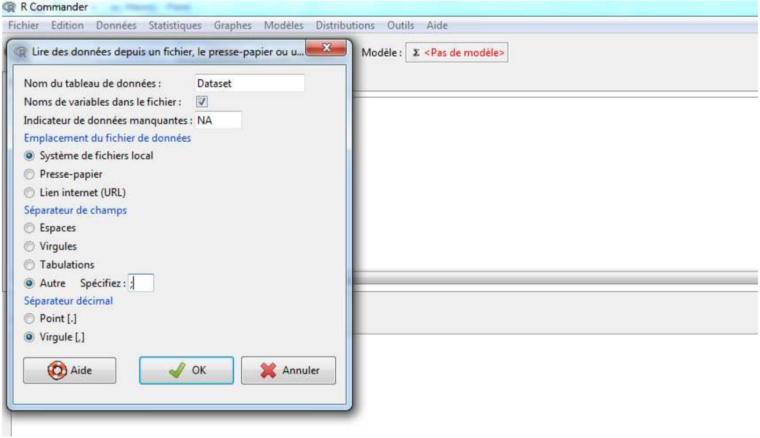
Nous n'avons pas pu mettre en évidence une différence significative entre les distributions des concentrations de médicaments A et B.

http://www.r-project.org/

Comparaison de 2 moyennes :
Exemple sous Rcmdr

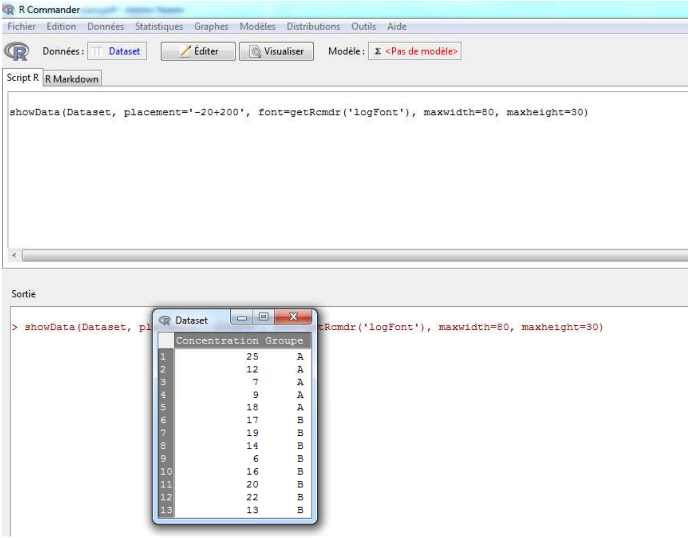


Comparaison de 2 moyennes :
Exemple sous Rcmdr



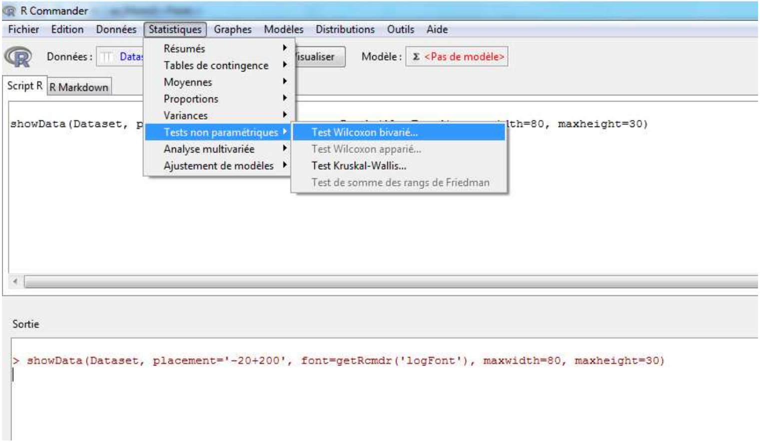
Comparaison de 2 moyennes :

Exemple sous Rcmdr



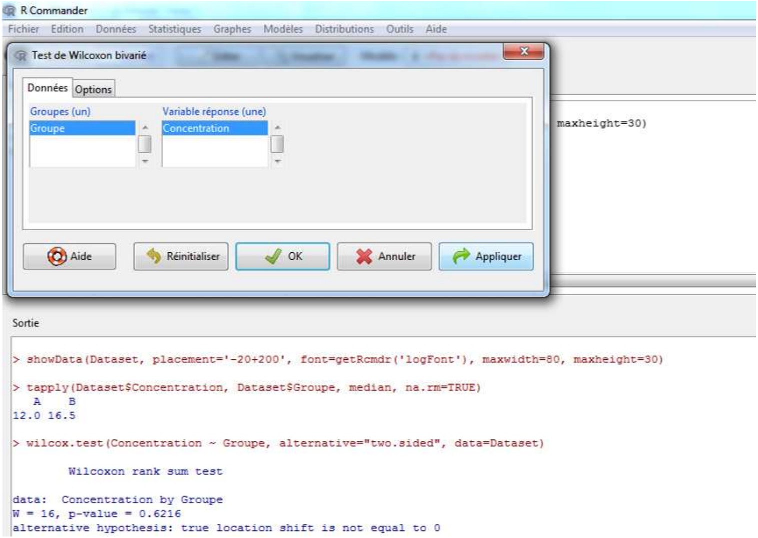
Comparaison de 2 moyennes :

Exemple sous Rcmdr



Comparaison de 2 moyennes :

Exemple sous Rcmdr



Plan

Introduction
Comparaison de moyennes
2 échantillons indépendants
+ de 2 échantillons indépendants
2 échantillons appariés
Comparaison de fréquences
2 échantillons indépendants
2 échantillons appariés
Corrélation linéaire
χ^2 d'indépendance
Conclusion

ANOVA à 1 facteur (1)

- Comparaison de 2 moyennes pour échantillons indépendants

  $\Rightarrow$ Test de Student, tests non paramétrique

- Comparaison de k moyennes (2 ou plus) simultanément pour échantillons indépendants

  $\Rightarrow$ Analyse de la Variance (ANOVA)

ANOVA à 1 facteur (2)

- Généralisation du test de Student
- Tester avec un seul test paramétrique si toutes les moyennes sont issues de la même population
- H_0 : les échantillons (groupes) proviennent de la même population

  $\iff$ Variabilité de l'ensemble uniquement conditionnée par les fluctuations d'échantillonnage
  $\iff H_0 : \mu_1 = \mu_2 = \ldots = \mu_k$
- H_1 : au moins 2 moyennes diffèrent entre elles

ANOVA à 1 facteur (3)

- Objectif de l'ANOVA

  - Décomposer la variabilité totale des données quantitatives (disposée en k groupes)

$\Rightarrow$ Différentes sources de variabilité :

$\rightarrow$ fluctuations individuelles (variabilité intra-groupe)
$\rightarrow$ fluctuations entre les groupes (variabilité inter-groupe)

- Idée générale

  Si variabilité inter-groupe > variabilité intra-groupe

$\Rightarrow$ 2 moyennes au moins $\neq$

$\Rightarrow$ Comparaison de variances

ANOVA à 1 facteur (4)

- Données observées :

Facteur	Groupe 1	Groupe 2	...	Groupe k
Effectif	n_1	n_2	...	n_k
Mesure	x_{11}	x_{12}	...	x_{1k}
Mesure	x_{21}	x_{22}	...	x_{2k}
Mesure	...	...	...	...
Mesure	$x_{n_1 1}$	$x_{n_2 2}$	...	$x_{n_k k}$
Moyennes	$\bar{x}_1$	$\bar{x}_2$	...	$\bar{x}_k$

ANOVA à 1 facteur (5)

- Limites

- Conditions de validité de l'ANOVA

  $\rightarrow$ critère d'intérêt (variable dépendante) quantitative
  $\rightarrow$ indépendance des observations
  $\rightarrow$ normalité du critère dans chaque groupe
  $\rightarrow$ homoscédasticité : homogénéité des variances

ANOVA à 1 facteur (6)

- Si conditions non vérifiées
  - ANOVA robuste à une certaine **hétéroscédasticité** (variance homogène)
  - ANOVA robuste face à une certaine **asymétrie** ou **applatissement** des distributions

$\longrightarrow$ Mais en cas de non respect majeur de la normalité

$\Rightarrow$ Transformation des données : $\log(x_{ij})$, $\sqrt{x_{ij}}$, $\frac{1}{x_{ij}}$, ...

$\Rightarrow$ ANOVA non paramétrique : test de Kruskall-Wallis

$\longrightarrow$ Mais en cas de non indépendance des résidus

$\Rightarrow$ changer de modèle

Principe (1)

- Comparaison de K populations (2 ou plus) simultanément pour échantillons indépendants

- Les fonctions de répartition $F_k(X)$ sont-elles toutes identiques ?

  - H_0 : $F_1(X) = ... = F_k(X) = ... = F_K(X)$
  - H_1 : au moins 1 distribution est différente des autres

Principe (2)

- Généralisation à K populations du test de la somme des rangs de Wilcoxon-Mann-Withney bilatéral

- Alternative non paramétrique à l'ANOVA quand v.a. non gaussienne

$\Rightarrow$ Statistique de test de Kruskal-Wallis :

$$H = \frac{12}{n(n+1)} \sum_{k=1}^{K} n_k (\bar{r}_k - \bar{r})^2$$

$\bar{r}$: moyenne globale des rangs
$\bar{r}_k$: moyenne des rangs pour le groupe k

Principe (3)

$\Rightarrow$ Expression d'une variabilité inter-classe

$\Leftrightarrow$ dispersion des moyennes conditionnelles autour de la moyenne globale

- $H \geq 0$
- si H_0 vérifiée, moyennes conditionnelles des rangs proches de la moyenne globale

  $\longleftrightarrow$ H proche de 0
- si H_1 vérifiée, H s'écarte de 0

Principe (4)

- si effectifs élevés ($n_k \geq 5$, $\forall\ k$), alors sous H_0, $H \sim \chi^2_{K-1ddl}$
- Région critique : $H \geq \chi^2_{K-1ddl}$
- Traitement des ex-aequo

$$\tilde{H} = \frac{H}{1 - \frac{\sum_{g=1}^{G}(t_g^3 - t_g)}{n^3 - n}}$$

G : nombre de valeurs distinctes ($G \leq n$)
t_g : nombre de valeurs pour la valeur $n°g$

Exemple (1)

- On souhaite comparer trois médicaments (A, B et C) contre les maux de têtes. La molécule C est aujourd'hui la référence sur le marché.

- 15 patients venants aux urgences pour des migraines sont traités par une des 3 molécules. Sur une échelle de 0 à 20, on demande aux patients d'évaluer leur douleur après que le médicament soit administré.

Exemple (2)

- Les résultats ont été rassemblés dans le tableau ci-dessous.

Molécule A	3	5	6	3	4
Molécule B	10	8	5	7	5
Molécule C	13	11	7	11	8

$\Rightarrow$ **Existe-t-il une différence d'efficacité entre les 3 molécules ?**

Exemple (3)

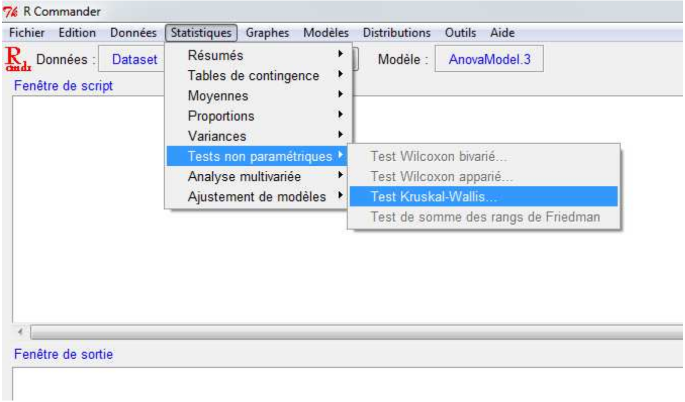
H_0 : "Il n'y a pas de différence entre les distributions de scores de douleur des 3 groupes."

$$\Leftrightarrow H_0 : F_A = F_B = F_C$$

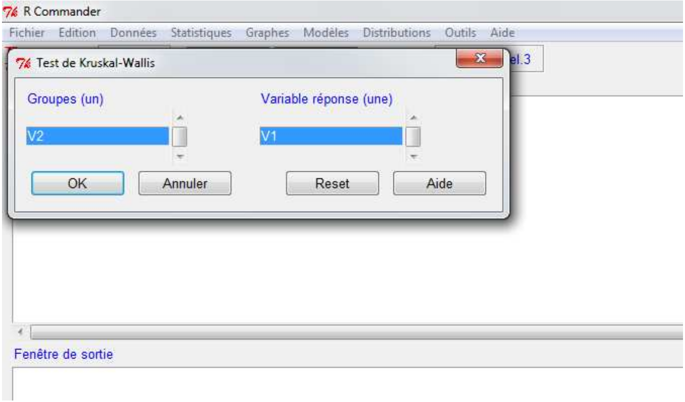
H_1 : "Il y a au moins un score dans l'un des groupes dont la distribution diffère de celle des 2 autres."

http://www.r-project.org/

Exemple (1) sous Rcmdr



Exemple (2) sous Rcmdr



Exemple (3) sous Rcmdr

Fenêtre de sortie

```
> tapply(Dataset$V1, Dataset$V2, median, na.rm=TRUE)
 1  2  3
 4  7 11

> kruskal.test(V1 ~ V2, data=Dataset)

        Kruskal-Wallis rank sum test

data:  V1 by V2
Kruskal-Wallis chi-squared = 9.3942, df = 2, p-value = 0.009122
```

$H = 9.39 > 5.99 = \chi^2_{2\ ddl}$ $(p_c < 0.05) \Rightarrow$ Rejet de H_0

Au seuil $\alpha = 0.05$, il semble qu'il y ait au moins deux molécules entrainant un score moyen de douleur qui soit différent.

Plan

L'échantillon A n'est pas indépendant de l'échantillon B

Echantillons dépendants (ou appariés, ou liés)

- observations obtenues chez les mêmes individus patient est son propre "témoin"

  ex : Etude avant/après (douleur avant et après la prise d'un antalgique)

  ex : Etude Cas/Témoin (deux pommades sur une même peau)

- observations obtenues chez des individus $\neq$ présentant des caractéristiques similaires

  ex : Etude Cas/Témoin (appariement sur les facteurs de risque)

  ex : Etude en cluster (deux souris d'une même litière, deux jumeaux, etc.)

Intérêt de l'appariement

- contrôler la variabilité inter-individuelle

  $\searrow$ variabilité inter-individuelle
  $\Rightarrow$ $\searrow$ variabilité totale ($\sigma \searrow$)
  $\Rightarrow$ augmenter la puissance statistique dans la comparaison

$\Rightarrow$ **Test statistique approprié** qui tient compte de l'appariement des observations

Contexte

- On observe 2 échantillons dépendants A et B de même taille $n = n_A = n_B$
- On définit le couple (X_A, X_B) pour chaque sujet/paire
- Définition de la variable aléatoire $D = X_A - X_B$: v.a. continue de moyenne μ_D.
- On observe un échantillon de taille n $\{d_1, d_2, ..., d_n\}$.

Remarque : cf. les tests de comparaison d'une moyenne observée à une valeur de référence.

$\Rightarrow$ Test usuel :

Comparaison de 2 moyennes
pour données appariées

Contexte

- On observe 2 échantillons dépendants A et B de même taille $n = n_A = n_B$
- On définit le couple (X_A, X_B) pour chaque sujet/paire
- Définition de la variable aléatoire $D = X_A - X_B$: v.a. continue de moyenne μ_D.
- On observe un échantillon de taille n $\{d_1, d_2, ..., d_n\}$.

Remarque : cf. les tests de comparaison d'une moyenne observée à une valeur de référence.

$\Rightarrow$ Test usuel : **Test paramétrique de STUDENT**

Limites

- Test nécessitant des effectifs de grandes tailles (n$\geq$30)
  $\rightarrow$ car théorème central limite valide

$\Rightarrow$ Solution :

Limites

- Test nécessitant des effectifs de grandes tailles (n$\geq$30)
  $\rightarrow$ car théorème central limite valide

$\Rightarrow$ Solution : TEST NON PARAMETRIQUE

Comparaison de 2 moyennes

Test de Wilcoxon pour données appariées

- Permet de comparer la distribution de 2 v.a. observées à partir de 2 échantillons appariés (A et B)

- Définition des v.a. (n sujets par groupe)
  - (X_A, X_B) : v.a. observées pour chaque paire
  - $X = X_A - X_B$: différence pour chaque paire

- Choix des hypothèses
  - H_0 : X_A et X_B ont la même distribution
  - H_1 : X_A et X_B n'ont pas la même distribution

Comparaison de 2 moyennes

Test de Wilcoxon pour données appariées

- Principe des rangs : classement des valeurs absolues $|X|$ en excluant les valeurs nulles et en notant le signe de la différence
- Soit k le nombre de différence non nulles
- Somme des rangs
  - $R(-)$: somme des rangs occupés par les différences négatives
  - $R(+)$: somme des rangs occupés par les différences positives
- si des ex-aequo entre des valeurs, on affecte la moyenne des rangs

Comparaison de 2 moyennes

Test de Wilcoxon pour données appariées

- Statistique de test :

$$W = min(R(+), R(-))$$

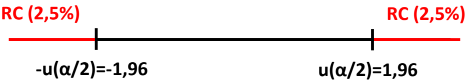
- Choix du seuil α (5%) et définition de la région critique

- Si $n > 20$, alors sous H_0 :

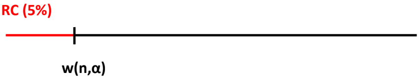
$$W \sim \mathcal{N}(n(n+1)/4, \sqrt{n(n+1)(2n+1)/24})$$

$$U = \frac{W - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}} \sim \mathcal{N}(0,1)$$

- test bilatéral, loi normale



- Si $n \leq 20$, lecture dans la table :



Comparaison de 2 moyennes

Test de Wilcoxon pour données appariées

Critical Values of the Wilcoxon Signed Ranks Test

n	Two-Tailed Test		One-Tailed Test	
	α = .05	α = .01	α = .05	α = .01
5	--	--	0	--
6	0	--	2	--
7	2	--	3	0
8	3	0	5	1
9	5	1	8	3
10	8	3	10	5
11	10	5	13	7
12	13	7	17	9
13	17	9	21	12
14	21	12	25	15
15	25	15	30	19
16	29	19	35	23
17	34	23	41	27
18	40	27	47	32
19	46	32	53	37
20	52	37	60	43
21	58	42	67	49
22	65	48	75	55
23	73	54	83	62
24	81	61	91	69
25	89	68	100	76
26	98	75	110	84
27	107	83	119	92
28	116	91	130	101
29	126	100	140	110
30	137	109	151	120

Comparaison de 2 moyennes

Test de Wilcoxon pour données appariées

- Si $u \in RC \rightarrow p < \alpha$:
  - Rejet de H_0 car moins de 5% de chances qu'elle soit vraie
  - Il semble que les 2 distributions soient différentes
- Si $u \notin RC \rightarrow p > \alpha$:
  - Non rejet de H_0 car plus de 5% de chances qu'elle soit vraie
  - On ne peut pas montrer que les 2 distributions soient différentes

Comparaison de 2 moyennes

Exemple pour données appariées

- Des chimistes ont mis au point une nouvelle méthode de mesure de concentration d'un toxique dans le sang. A partir de 12 prélèvements sanguins divisés en deux tubes, la nouvelle méthode (A) est comparée avec la méthode existante (B).

- Les concentrations obtenues sont présentées dans le tableau suivant :

Numéro tube	1	2	3	4	5	6	7	8	9	10	11	12
Méthode A	55	25	21	7	16	14	7	9	5	4	4	1
Méthode B	17	17	14	11	10	9	8	5	3	2	1	0

X_A et X_B : concentrations de toxiques sur les 12 prélèvements obtenus respectivement à l'aide des méthodes A et B
$X = X_A - X_B$: différence de chacune des paires.

On cherche à répondre à la question suivante :
"Les concentrations de toxique obtenues avec la nouvelle méthode sont telles significativement différente de celles obtenues avec la méthode de référence ?" (risque d'erreur de 1ère espèce fixé à 5%) ?

- Quel test statistique doit-on réaliser ?

Comparaison de 2 moyennes

Exemple pour données appariées

On cherche à répondre à la question suivante :
"Les concentrations de toxique obtenues avec la nouvelle méthode sont telles significativement différente de celles obtenues avec la méthode de référence ?" (risque d'erreur de 1ère espèce fixé à 5%) ?

- Quel test statistique doit-on réaliser ?

Test de Wilcoxon pour données appariées

Conditions $\Rightarrow$ échantillons de petites tailles ($n < 30$)

Test d'une $\neq$ $\Rightarrow$ test bilatéral

Comparaison de 2 moyennes

Exemple pour données appariées

- Quelle est l'hypothèse nulle H_0 de ce test ?
- Quelle est l'hypothèse alternative H_1 de ce test ?

- Quelle est l'hypothèse nulle H_0 de ce test ?

- Quelle est l'hypothèse alternative H_1 de ce test ?

H_0 : les distributions issues des 2 méthodes de prélèvements sont identiques, autrement dit X_A et X_B ont la même distribution

H_1 : les distributions issues des 2 méthodes de prélèvements ne sont pas identiques, c'est-à-dire que X_A et X_B n'ont pas la même distribution

Comparaison de 2 moyennes

Exemple pour données appariées

On travaille sur les valeurs absolues des rangs
Remarque : on exclue les valeurs nulles

Numéro tube	1	2	3	4	5	6	7	8	9	10	11	12
$\delta_{\text{Méthode A - Méthode B}}$	38	8	7	-4	6	5	-1	4	2	2	3	1
Rang	12	11	10	6.5	9	8	1.5	6.5	3	4	5	1.5

Rang négatif	1.5	6.5								
Rang positif	12	11	10	9	8	6.5	3	4	5	1.5

La statistique de test est définie de la manière suivante :

$$W = min(R(-), R(+))$$

$R(-)$: $\sum$ des rangs occupées par les différences négatives
$R(+)$: $\sum$ des rangs occupées par les différences positives

En pratique, on a donc :

$$R_{-} = 1.5 + 6.5 = 8$$

$$R_{+} = 12 + 11 + 10 + 9 + 8 + 6.5 + 3 + 4 + 5 + 1.5 = 70$$

$$\Rightarrow W = 8$$

Comparaison de 2 moyennes

Exemple pour données appariées

$n = 12 < 20 \Rightarrow$ on lit dans la table la valeur critique définissant la région critique.

Au seuil $\alpha = 0.05$, pour $n = 12$, on a pour valeur critique la valeur 13 (pour un test bilatéral)

$W \in$ à la région critique

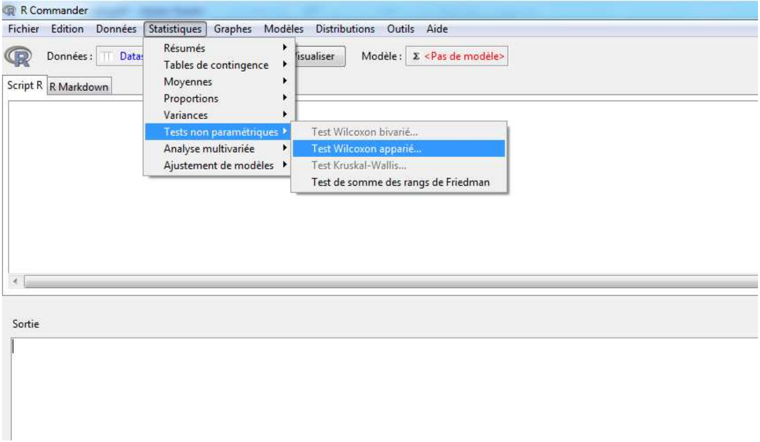
$\Rightarrow$ Rejet H_0 au risque 5% de se tromper

Il semble donc y avoir une différence significative entre les distributions des concentrations de toxiques issues de la méthode A et de la méthode B.

http://www.r-project.org/

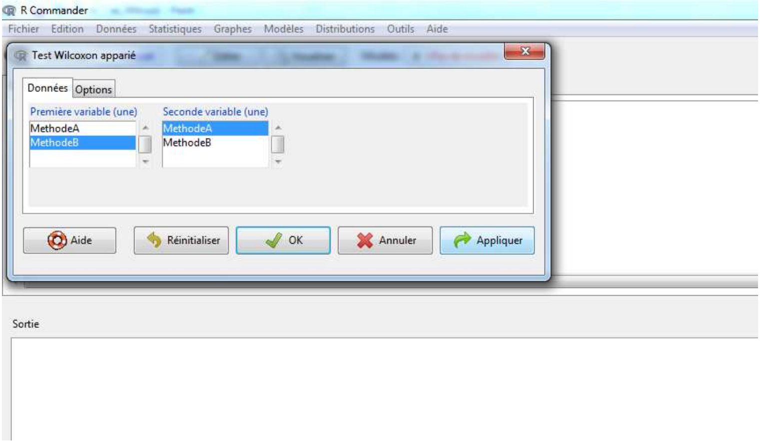
Comparaison de 2 moyennes

Exemple pour données appariées sous Rcmdr



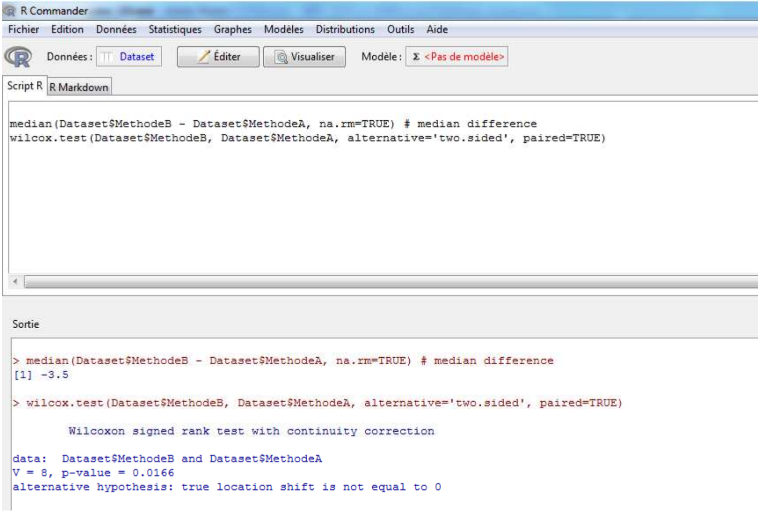
Comparaison de 2 moyennes

Exemple pour données appariées sous Rcmdr



Comparaison de 2 moyennes

Exemple pour données appariées sous Rcmdr



Plan

Plan

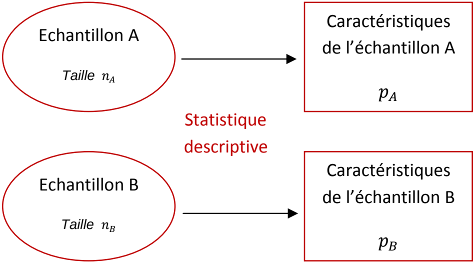
Test paramétrique de Student

pour la comparaison de 2 fréquences observées
(expression incorrecte ! !)

$\Longleftrightarrow$ pour la comparaison d'une fréquence théorique issue d'une population représentée par un échantillon observé à une fréquence théorique issue d'une autre population représentée par un autre échantillon

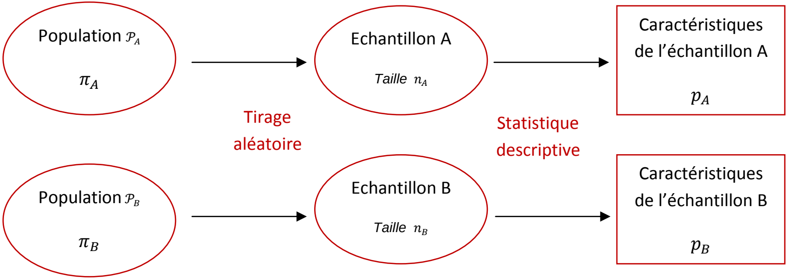
Comparaison de 2 fréquences :

Test paramétrique de Student



Comparaison de 2 fréquences :
Test paramétrique de Student

avec $n_A > 30$ et $n_B > 30$
et $n_A p_A > 5$, $n_A(1 - p_A) > 5$, $n_B p_B > 5$, $n_B(1 - p_B) > 5$



$$H_0 : \pi_A = \pi_B = \pi$$

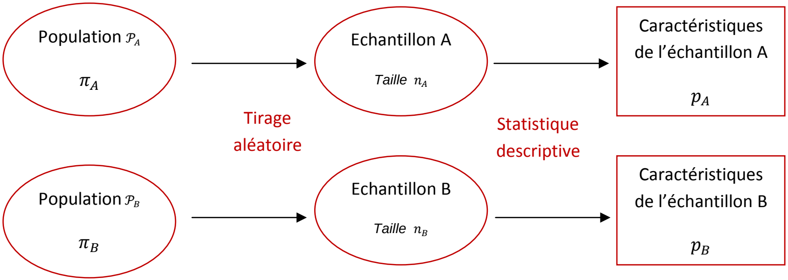
Conclusion

- non rejet de H_0 : la différence entre π_A et π_B semble due aux fluctuations d'échantillonnage,
  i.e. échantillons A et B semblent représentatifs de la population P
- rejet de H_0 : la différence entre π_A et π_B n'est pas due au hasard,
  i.e. échantillons A et B semblent représentatifs de 2 populations distinctes

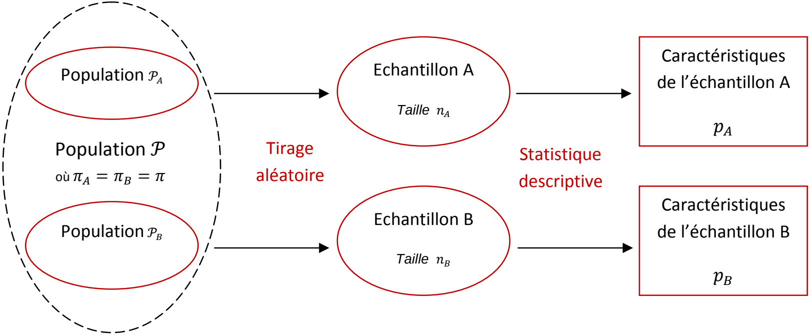
$\Rightarrow$ attention : il existe **toujours** une probabilité de se tromper

α : probabilité de se tromper en rejetant H_0
β : probabilité de se tromper en ne rejetant pas H_0, i.e. en rejetant H_1

Non rejet de H_0 au risque de 1ère espèce α



Rejet de H_0 au risque de 1ère espèce α



Comparaison de 2 fréquences :

Test paramétrique de Student

- Philosophie similaire à celle d'un test paramétrique de Student pour comparaison de 2 moyennes
- Non présenté car Test du χ^2 plus intéressant !
- Problème quand les effectifs sont petits
  - Aucune hypothèse possible
  - Solution : calcul exact de la probabilité critique

Comparaison de 2 fréquences :

Test exact de Fisher

- Adapté aux petits échantillons
- Principe : calculer exactement la probabilité critique p (probabilité d'obtenir des échantillons au moins aussi éloignés de H_0 que celui qui est observé)
- Mise en oeuvre et raisonnement repose sur la **loi hypergéométrique**
- Peut nécessiter des combinaisons de calculs très nombreuses en fonction du nombre de modalités comparées
  $\Rightarrow$ 2 modalités (succès/échec)

Loi hypergéométrique

Epreuve aléatoire de Bernouilli

→ S (succès) : $\Pr(S) = p, 0 \leq p \leq 1$

→ E (échec) : $\Pr(E) = 1 - p = q, 0 \leq q \leq 1$

Si n répétitions **non indépendantes** de l'épreuve,

et R : v.a. caractérisant le nombre de succès,

R ~ H(N, n, p)

- N : Taille de la population
- n : Nb de répétitions
- p : Probabilité de succès

Exemple :

-tirage sans remise dans un lot de taille finie N

➔ Contenu de la population est modifié à chaque tirage → il n'y a plus indépendance

Comparaison de 2 fréquences :
Test exact de Fisher

On observe :

	Modalité 1	Modalité 2	Total
Groupe 1	a	b	a+b
Groupe 2	c	d	c+d
Total	a+c	b+d	N

Probabilité d'obtenir cette configuration : *p*

$$p(obs) = \frac{C_{a+c}^{a} C_{b+d}^{b}}{C_{N}^{a+b}} = \frac{(a+b)!(c+d)!(a+c)!(b+d)!}{a!b!c!d!N!}$$

Comparaison de 2 fréquences :
Test exact de Fisher

- De la même manière, on calcule toute les probabilités p_i pour tous les échantillons i encore plus éloignés de H_0 que l'échantillon observé,
- c'est-à-dire les tableaux composés des effectifs a', b', c', d' tels que

$$\left| \frac{a'}{a'+c'} - \frac{b'}{b'+d'} \right| \geq \left| \frac{a}{a+c} - \frac{b}{b+d} \right|$$

- Calcul de la probabilité critique :

$$p = p(obs) + \sum_i p_i$$

Un peu de discussion...

Attention à la surinterprétation des résultats...
surtout dans le cas de petits échantillons

$\Rightarrow$ le risque α existe et $\alpha \neq 0$

Conclusion statistique

- Si $p_c < \alpha$.
  - Rejet de H_0 car moins de α % de chance qu'elle soit vraie.
- Si $p_c > \alpha$.
  - Non rejet de H_0 car plus de α % de chance qu'elle soit vraie.

Conclusion épidémiologique

- Si Rejet de H_0,
  - cas bilatéral : Il semble que les 2 proportions théoriques π_A et π_B soient significativement différente.
  - cas unilatéral : Il semble que la proportion théorique π_A soit significativement inférieure (supérieure) à la proportion théorique π_B
- Si Non rejet de H_0
  - cas bilatéral et unilatéral : On ne peut pas exclure le fait que les 2 échantillons soient issus de la même population.

Cas unilatéral

$\rightarrow$ stratégie identique mais on ne compte que les échantillons i encore plus éloignés de H_0 mais que d'un seul côté

Comparaison de 2 fréquences :

Exemple

- Comparaison de 2 traitements A et B contre infection nosocomiale
  $\Rightarrow$ proportions de patients infectés

- Hypothèses :
  $H_0 : \pi_A = \pi_B = \pi$
  $H_1 : \pi_A \neq \pi_B$

- Etude pilote : N=16 patients
  $\rightarrow$ 7 patients dans le groupe A
  $\rightarrow$ 9 patients dans le groupe B

Comparaison de 2 fréquences :
Exemple

On observe :

	Infectés	Non infectés	Total
Groupe A	6	1	7
Groupe B	2	7	9
Total	8	8	16

Idée

Calculer sous H_0 la probabilité (exacte) d'obtenir un écart entre les groupes $\geq$ à celui observé en conservant les mêmes totaux (lignes et colonnes)

$\Rightarrow$ Probabilité critique de Fisher : somme de ces probabilités

Comparaison de 2 fréquences :

Exemple

On observe :

	Infectés	Non infectés	Total
Groupe A	6	1	7
Groupe B	2	7	9
Total	8	8	16

1 seule autre possibilité avec les mêmes totaux

	Infectés	Non infectés	Total
Groupe A	7	0	7
Groupe B	1	8	9
Total	8	8	16

Comparaison de 2 fréquences :
Exemple

On observe :

	Infectés	Non infectés	Total
Groupe A	6	1	7
Groupe B	2	7	9
Total	8	8	16

Probabilité d'obtenir cette configuration : p_1

$$p_1 = \frac{7!9!8!8!}{6!1!2!7!16!}$$

Comparaison de 2 fréquences :
Exemple

Seconde configuration avec même totaux mais des écarts entre les groupes supérieurs à ce qu'on observe

	Infectés	Non infectés	Total
Groupe A	7	0	7
Groupe B	1	8	9
Total	8	8	16

Probabilité d'obtenir cette configuration : p_2

$$p_2 = \frac{7!9!8!8!}{7!0!1!8!16!}$$

Comparaison de 2 fréquences :
Exemple

$\Rightarrow$ Probabilité d'avoir l'une ou l'autre des configurations possibles :

$$p = p_1 + p_2 = \frac{7!9!8!8!}{6!1!2!7!16!} + \frac{7!9!8!8!}{7!0!1!8!16!} \approx 0.0203$$

- ici, hypothèse **bilatérale** $\Rightarrow$ écart doit être envisagé dans les 2 sens $\Rightarrow$ probabilité critique de Fisher

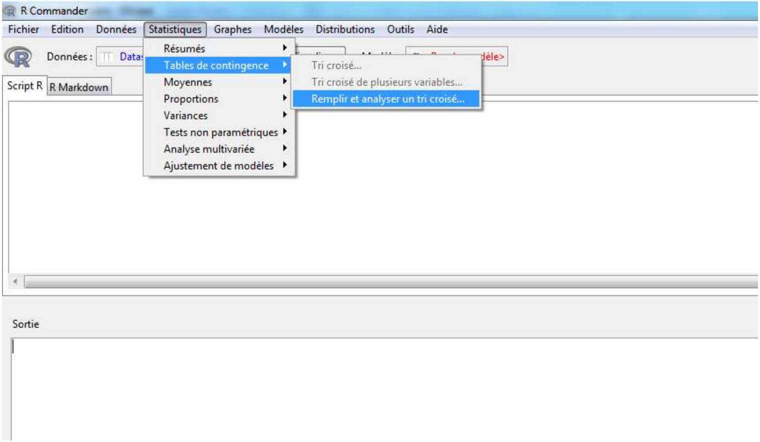
$$p_c = 2 * p = 0.0406 < 0.05$$

$\Rightarrow$ Rejet de H_0 au seuil $\alpha = 5\%$

`http://www.r-project.org/`

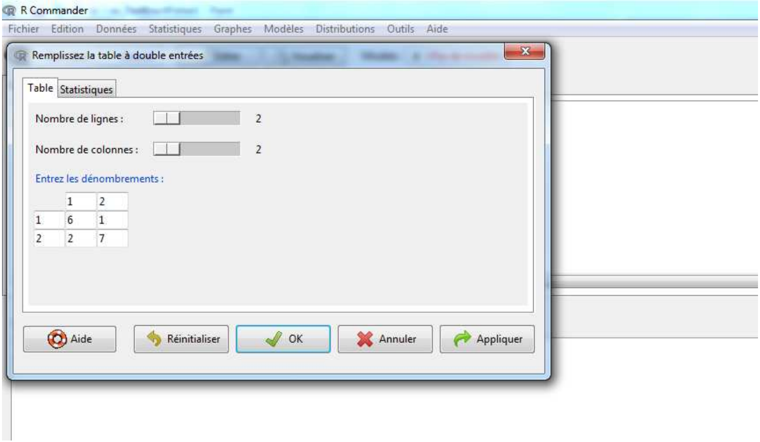
Comparaison de 2 fréquences

Exemple sous Rcmdr



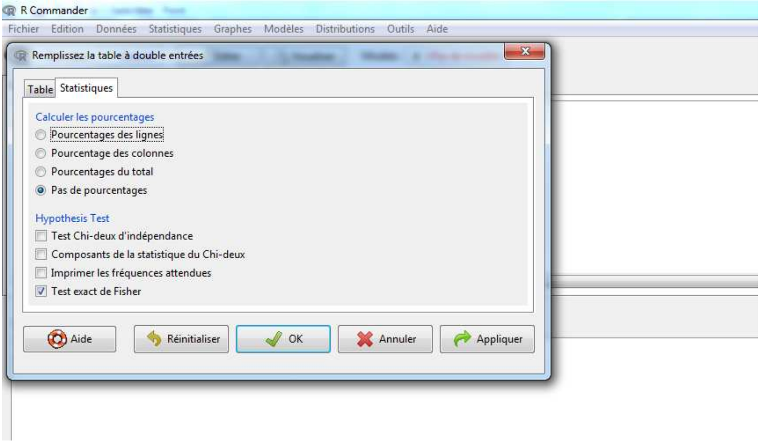
Comparaison de 2 fréquences

Exemple sous Rcmdr



Comparaison de 2 fréquences

Exemple sous Rcmdr



Comparaison de 2 fréquences

Exemple sous Rcmdr

R Commander

Fichier Edition Données Statistiques Graphes Modèles Distributions Outils Aide

Données : <Pas de données> Éditer Visualiser Modèle : Σ <Pas de modèle>

Script R | R Markdown

```
.Table <- matrix(c(6,1,2,7), 2, 2, byrow=TRUE)
rownames(.Table) <- c('1', '2')
colnames(.Table) <- c('1', '2')
.Table  # Counts
.Test <- chisq.test(.Table, correct=FALSE)
.Test
remove(.Test)
remove(.Table)
```

Sortie

```
> .Table <- matrix(c(6,1,2,7), 2, 2, byrow=TRUE)

> rownames(.Table) <- c('1', '2')

> colnames(.Table) <- c('1', '2')

> .Table  # Counts
  1 2
1 6 1
2 2 7

> .Test <- chisq.test(.Table, correct=FALSE)

> .Test

        Pearson's Chi-squared test

data:  .Table
X-squared = 6.3492, df = 1, p-value = 0.01174
```

Plan

L'échantillon A n'est pas indépendant de l'échantillon B

Echantillons dépendants (ou appariés, ou liés)

- observations obtenues chez les mêmes individus patient est son propre "témoin"

  ex : Etude avant/après (douleur avant et après la prise d'un antalgique)

  ex : Etude Cas/Témoin (deux pommades sur une même peau)

- observations obtenues chez des individus $\neq$ présentant des caractéristiques similaires

  ex : Etude Cas/Témoin (appariement sur les facteurs de risque)

  ex : Etude en cluster (deux souris d'une même litière, deux jumeaux, etc.)

Comparaison de 2 fréquences

Test de Mc Nemar

- Exemple classique : essai thérapeutique en "cross-over"
  $\Rightarrow$ patient est son propre "témoin"
  traitements comparés administrés successivement (*ordre d'administration tiré au sort*)

- Critères d'intérêt : Succès(+) vs. Echec (-)

		Avant traitement	Après traitement
Sujets	1	+	-
	2	-	+
	3	+	+
	4	-	-
	...	...	...
	N	+	+

Comparaison de 2 fréquences
Test de Mc Nemar

On réalise un test de Mc Nemar pour comparer 2 proportions théoriques issues de 2 populations représentées par 2 échantillons

$\Rightarrow$ 4 configurations possibles :

+	+	q sujets
+	-	r sujets
-	+	s sujets
-	-	t sujets

		Après +	Après -	
Avant	+	q	r	q+r
	-	s	t	s+t
		q+s	r+t	N=q+r+s+t

Comparaison de 2 fréquences

Test de Mc Nemar

- Problème posé : Comparer proportions

$p_1(+)$ avant : $\frac{q+r}{N} = p_1$

et

$p_2(+)$ après : $\frac{q+s}{N} = p_2$

$$\Delta = p_1 - p_2 = \frac{q+r}{N} - \frac{q+s}{N} = \frac{r}{N} - \frac{s}{N}$$

$\Rightarrow$ Seuls interviennent les effectifs r et s relatifs à un changement

Comparaison de 2 fréquences

Test de Mc Nemar

- Remarques

  $\Rightarrow$ La différence entre les proportions observées de succès dépend

  - Effectifs de paires discordantes
  - Effectif total

- Sous H_0, on s'attend à avoir la même proportion de paires discordantes ($p_{(+-)}$ et $p_{(-+)}$)

- Soit $\pi_{(+-)}$ et $\pi_{(-+)}$: proportions théoriques de paires discordantes

$$H_0 : \pi_{(+-)} = \pi_{(-+)} = 0.5$$
$$H_1 : \pi_{(+-)} \neq \pi_{(-+)} \quad \text{(en bilatéral)}$$

$\Rightarrow$ Comparaison d'une proportion observée ("+-" ou "-+" parmi total de paires discordantes) à une proportion théorique (=0.5)

Comparaison de 2 fréquences

Test de Mc Nemar

$$H_0 : \pi_{(+-)} = \pi_{(-+)} = 0.5$$
$$H_1 : \pi_{(+-)} \neq \pi_{(-+)} \quad \text{(en bilatéral)}$$

- Statistique

  $\Rightarrow$ R=nombre de sujets (+-) parmi les $n = r + s$ ayant expérimenté un changement

  $\Rightarrow$ Si H_0 est vraie : $R \sim B(n, \pi_{(+-)} = 0.5)$

$$E(R) = n\pi_{(+-)} = n/2$$
$$Var(R) = n\pi_{(+-)}(1 - \pi_{(+-)}) = n/4$$

- si $n > 30$, approximation de la loi binomiale par une loi normale

  Sous H_0,

$$R \sim N\left(\mu = \frac{n}{2}, \sigma = \frac{\sqrt{n}}{2}\right)$$

$$\Rightarrow T = \frac{R - \frac{n}{2}}{\frac{\sqrt{n}}{2}} \sim N(0,1)$$

- Réalisation pratique

$$|t_{exp}| = \left|\frac{r - \frac{n}{2}}{\frac{\sqrt{n}}{2}}\right| = \left|\frac{r-s}{\sqrt{r+s}}\right|$$

$\Rightarrow$ $|t_{exp}| < t_{table,\alpha} \Rightarrow$ non Rejet de H_0 au seuil α

$\Rightarrow$ $|t_{exp}| \geq t_{table,\alpha} \Rightarrow$ Rejet de H_0 au seuil α

Comparaison de 2 fréquences

Test de Mc Nemar

- Remarque :

Si $T = \frac{R - \frac{n}{2}}{\frac{\sqrt{n}}{2}} \sim N(0,1)$, alors

$$T^2 = \frac{(R-S)^2}{R+S} \sim \chi^2_{1ddl}$$

Comparaison de 2 fréquences

Exemple pour données appariées

Contexte

Dans un service de neurologie, une étude est mise en place dans le cadre de douleurs polyneuropathiques : elle vise à comparer le traitement de référence par tramadol à un traitement nouvellement mis sur le marché. On s'intéresse à l'amélioration de la satisfaction des patients vis à vis de la douleur ressentie.

On interroge 60 patients pris en charge pour une polyneuropathie avant la prise du nouveau médicament lorsqu'il sont traités par tramadol et après l'introduction du nouveau médicament (+ : douleur ; - : pas de douleur).

On observe :

		Après	
		+	-
Avant	+	18	26
	-	11	5

Etape 1 : Choix des hypothèses

question d'intérêt :
le choix de la nouvelle thérapeutique semble-t-il pertinent ?

$\Leftrightarrow$ la proportion de patients ayant des douleurs polyneuropathiques après la prise du nouveau médicament est-elle moindre que celle avant ?

$\Rightarrow$ test unilatéral

$$H_0 : \pi_{(+-)} = \pi_{(-+)} = 0.5$$

$$H_1 : \pi_{(+-)} > \pi_{(-+)}$$

Comparaison de 2 fréquences

Exemple pour données appariées

Etape 2 : Choix du test, statistique de test et loi sous H_0

Sous H_0, comme $n = 26 + 11 = 37 > 30$, on a :

$$T = \frac{R - \frac{n}{2}}{\frac{\sqrt{n}}{2}} \sim N(0,1)$$

R : v.a. représentant le nombre de gens qui ont changé d'avis dans un sens bien précis (par ex., de + vers -)
S : v.a. telle que $S = n - R$ (nombre de gens qui ont changé d'avis dans l'autre sens.
On montre que

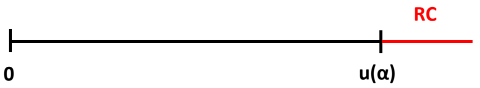
$$T = \frac{R - (R+S)/2}{\sqrt{R+S}/2} = \frac{\frac{2R-R-S}{2}}{\sqrt{R+S}/2} = \frac{R-S}{\sqrt{R+S}}$$

Comparaison de 2 fréquences

Exemple pour données appariées

Etape 3 : Choix du niveau de signification α, Définition de la région critique (RC)

ICI : test unilatéral



en pratique, $\alpha = 0.05$
$\Rightarrow$ RC : $[1.64; +\infty[$

Etape 4 : Calcul de statistique de test

Analyse numérique réalisée à l'aide des observations

Il y a deux manières de calculer t_{exp}

$$t_{exp} = \frac{r - n/2}{\sqrt{n}/2} = \frac{r - s}{\sqrt{r+s}} = \frac{26 - 37/2}{\sqrt{37}/2} = \frac{26 - 11}{\sqrt{26+11}} \approx 2.47$$

Etape 5 : Conclusion statistique

- $t_{exp} \approx 2.47 \in [1.64; +\infty[$

$\Rightarrow$ $t_{exp} \in RC \rightarrow p_c < \alpha$

$\rightarrow$ Rejet de H_0 au seuil $\alpha = 5\%$

Etape 6 : Conclusion épidémiologique

$\Rightarrow$ Il semble que la proportion de patients sans douleur après la prise du nouveau traitement soit moindre que celle des patients avec des douleurs neuropathiques. Je peux affirmer cela avec moins de 5% de chances de me tromper

$\Rightarrow$ La nouvelle thérapeutique semble être intéressante dans l'amélioration de la satisfaction des patients vis à vis de leur prise en charge.

- si $n < 30$, test binomial basé sur le calcul de la probabilité (p_c) sous H_0

Exemple (petits échantillons)

On étudie un échantillon de 30 patients
On observe

		Après a-Tab +	Après a-Tab -
Avant a-Tab	+	10	12
	-	5	3

Comparaison de 2 fréquences

Exemple pour données appariées

Etape 1 : Choix des hypothèses

question d'intérêt :
le choix de la nouvelle thérapeutique semble-t-il pertinent ?

$\Leftrightarrow$ la proportion de patients ayant des douleurs polyneuropathiques après la prise du nouveau médicament est-elle moindre que celle avant ?

$\Rightarrow$ test unilatéral

$$H_0 : \pi_{(+-)} = \pi_{(-+)} = 0.5$$

$$H_1 : \pi_{(-+)} > \pi_{(+-)}$$

Comparaison de 2 fréquences

Exemple pour données appariées

Etape 2 : Choix du test, statistique de test et loi sous H_0

Sous H_0, comme $n = 12 + 5 = 17 < 30$, on a :

$$S \sim B(n = 17, \pi_{(+-)} = 0.5)$$

où S=nombre de sujets (-+) parmi les $n = r + s$ ayant expérimenté un changement

Etape 3 : Choix du niveau de signification α

en pratique, $\alpha = 0.05$

Etape 4 : Calcul de la probabilité critique

$$
\begin{aligned}
p_c &= Pr(S \geq 12) = 1 - Pr(S < 12) \\
&= 1 - Pr(S = 0) - Pr(S = 1) - \ldots - Pr(S = 11) \\
&= 1 - C_{17}^{0} 0.5^{0} (1 - 0.5)^{17} - C_{17}^{1} 0.5^{1} (1 - 0.5)^{16} \\
&- \ldots - C_{17}^{11} 0.5^{11} (1 - 0.5)^{6} \\
&\approx 0.071
\end{aligned}
$$

Comparaison de 2 fréquences

Exemple pour données appariées

Etape 5 : Conclusion statistique

- $p_c \approx 0.071 > \alpha = 0.05$

  $\rightarrow$ Non Rejet de H_0 au seuil $\alpha = 5\%$

Etape 6 : Conclusion épidémiologique

$\Rightarrow$ Il semble que la proportion de patients sans douleur après la prise du nouveau traitement ne soit pas significativement différente de celle des patients avec des douleurs neuropathiques. Je peux affirmer cela avec moins de 5% de chances de me tromper

$\Rightarrow$ notre étude ne nous a pas permis de mettre en évidence une efficacité du nouveau traitement

Plan

Introduction

- 2 v.a. quantitatives X et Y sont mesurées sur les mêmes sujets et sont supposées liées

$\Rightarrow$ **quel est l'intensité de ce lien ?**

$\Rightarrow$ **existe-t-il un lien (une association) entre ces 2 v.a. ?**

Introduction

- 2 v.a. quantitatives X et Y sont mesurées sur les mêmes sujets et sont supposées liées

$\Rightarrow$ **quel est l'intensité de ce lien ?**

$\Leftrightarrow$ quelle est la valeur du coefficient de corrélation ?

$\Rightarrow$ **existe-t-il un lien (une association) entre ces 2 v.a. ?**

Introduction

- 2 v.a. quantitatives X et Y sont mesurées sur les mêmes sujets et sont supposées liées

$\Rightarrow$ **quel est l'intensité de ce lien ?**

$\Leftrightarrow$ quelle est la valeur du coefficient de corrélation ?

$\Rightarrow$ **existe-t-il un lien (une association) entre ces 2 v.a. ?**

$\Leftrightarrow$ le coefficient de corrélation des 2 v.a. est-il significativement $\neq$ de 0 ?

Exemples

- lien entre “Poids” et “Taille”
- lien entre ”Alimentation” et “Prise de poids”
- lien entre ”Âge” et “Déclin cognitif”
- lien entre ”Âge” et “Taux de cholestérol”
- lien entre ”Âge” et “Décès”

Remarque :

pour l'étude d'1 lien entre le statut tabagique et la survenue d'un cancer du poumon, que faire ?

Exemples

- lien entre “Poids” et “Taille”
- lien entre ”Alimentation” et “Prise de poids”
- lien entre ”Âge” et “Déclin cognitif”
- lien entre ”Âge” et “Taux de cholestérol”
- lien entre ”Âge” et “Décès”

Remarque :

pour l'étude d'1 lien entre le statut tabagique et la survenue d'un cancer du poumon, que faire ?

$\rightarrow$ analyse de la relation entre 2 variables qualitatives

Exemples

- lien entre “Poids” et “Taille”
- lien entre ”Alimentation” et “Prise de poids”
- lien entre ”Âge” et “Déclin cognitif”
- lien entre ”Âge” et “Taux de cholestérol”
- lien entre ”Âge” et “Décès”

Remarque :

pour l'étude d'1 lien entre le statut tabagique et la survenue d'un cancer du poumon, que faire ?

$\rightarrow$ analyse de la relation entre 2 variables qualitatives

$\Rightarrow$ Test de chi-deux, Test de comparaison de deux fréquences + calcul OR

Coefficient de corrélation de Pearson

- Soient 2 v.a. quantitatives continues X et Y (avec une distribution jointe supposée bi-normale), de moyennes μ_X et μ_Y et de variances σ_X^2 et σ_Y^2

  $\rightarrow$ En pratique, bi-normalité pas évidente, déjà bien si X et Y sont considérées normales

- Echantillon composé de n individus ($i = 1, ..., n$)

$\Rightarrow$ On observe n couples (y_i, x_i)

Coefficient de corrélation de Pearson

- Coefficient de corrélation entre X et Y :

$$\rho = \frac{Cov(X, Y)}{\sqrt{\sigma_X^2 \sigma_Y^2}}$$

$\Rightarrow$ ρ mesure l'intensité de la relation entre deux variables quantitatives X et Y si cette relation est linéaire

Coefficient de corrélation de Spearman

- Limites du coefficient de corrélation de Pearson

$\Rightarrow$ Bi-normalité des 2 v.a. X et Y pas vérifiée

$\rightarrow$ Solutions :

Coefficient de corrélation de Spearman

- Limites du coefficient de corrélation de Pearson

$\Rightarrow$ Bi-normalité des 2 v.a. X et Y pas vérifiée

$\rightarrow$ Solutions :

- utilisation de transformée de X et Y pour se rapprocher de la bi-normalité
  (ou tout du moins de la normalité de chacune des v.a.)
- utiliser un coefficient basé sur les rangs
  et un test non-paramétrique

$\Rightarrow$ corrélation de Spearman

Coefficient de corrélation de Spearman

- Soient 2 v.a. quantitatives continues X et Y, de moyennes μ_X et μ_Y et de variances σ^2_X et σ^2_Y
- Echantillon composé de n individus ($i = 1, ..., n$)

$\Rightarrow$ On observe n couples (y_i, x_i)

$\rightarrow$ pour X et Y, on affecte le rang correspondant à chacune des valeurs

$\Rightarrow$ On observe n couples de rang (R_{Yi}, R_{Xi})

Coefficient de corrélation de Spearman

- Soient 2 v.a. quantitatives continues X et Y, de moyennes μ_X et μ_Y et de variances σ_X^2 et σ_Y^2
- Echantillon composé de n individus ($i = 1, ..., n$)

$\Rightarrow$ On observe n couples (y_i, x_i)

$\rightarrow$ pour X et Y, on affecte le rang correspondant à chacune des valeurs

$\Rightarrow$ On observe n couples de rang (R_{Yi}, R_{Xi})

$\Rightarrow$ Coefficient de corrélation entre X et Y :

$$\rho = \frac{Cov(R_Y, R_X)}{\sqrt{\sigma^2_{R_X}\sigma^2_{R_Y}}}$$

Coefficient de corrélation de Spearman

- Coefficient de corrélation entre X et Y :

$\rightarrow$ + simple de s'intéresser à $D = R_{Xi} - R_{Yi}$, on alors :

$$\rho = 1 - \frac{6\sum_{i=1}^{n} D_i^2}{n * (n^2 - 1)}$$

Interprétation (1)

- Interprétation du coefficient de corrélation de Spearman :

$\Rightarrow$ on montre facilement que $\rho \in [-1; +1]$

$\rightarrow$ $r = 1$: lien linéaire parfait dans le même sens

$\rightarrow$ $r = -1$: lien linéaire parfait dans le sens inverse

$\rightarrow$ $|r| > 0.5$: lien linéaire fort

$\rightarrow$ $0.3 < |r| < 0.5$: lien linéaire moyen

$\rightarrow$ $0.1 < |r| < 0.3$: lien linéaire faible

$\rightarrow$ $r = 0$: pas de liaison linéaire

$\Rightarrow$ identique à celle du coefficient de corrélation de Pearson

Interprétation (2)

- Interprétation du coefficient de corrélation linéaire :

$\rightarrow$ si X et Y sont 2 v.a. indépendantes alors

$$\rho = 0$$

En effet, si X et Y sont 2 v.a. indépendantes, alors $E(X.Y) = E(X).E(Y)$

$\rightarrow Cov(X, Y) = E(X.Y) - E(X).E(Y) = 0$

$$\Rightarrow \rho = 0$$

$\rightarrow$ si $\rho = 0$ et X et Y normalement distribuées $\Rightarrow$ X et Y sont 2 v.a. indépendantes

Exemple

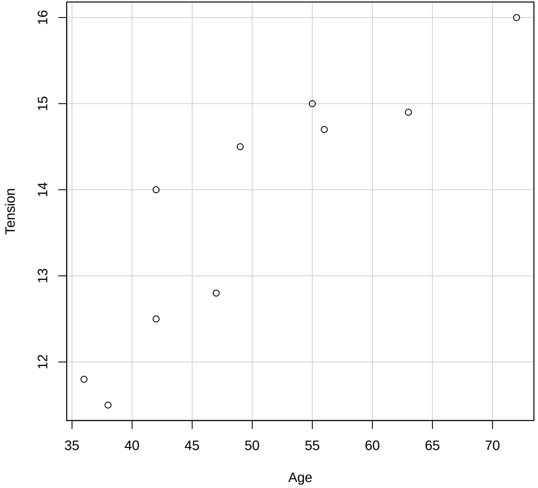
- Soient Y et X v.a. représentant respectivement la tension artérielle systolique et l'âge de femmes inclues dans un échantillon

- On réalise un échantillon de 10 femmes, les résultats sont présentés dans le tableau suivant :

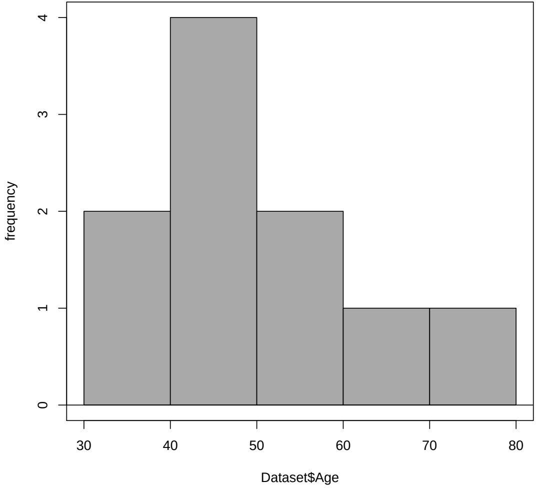
Sujet	1	2	3	4	5	6	7	8	9	10
x (ans)	56	42	72	36	63	47	55	49	38	42
y (mmHg)	14.7	12.5	16.0	11.8	14.9	12.8	15.0	14.5	11.5	14.0

$\Rightarrow$ On souhaite tester si la tension artérielle systolique diffère avec l'âge chez les femmes

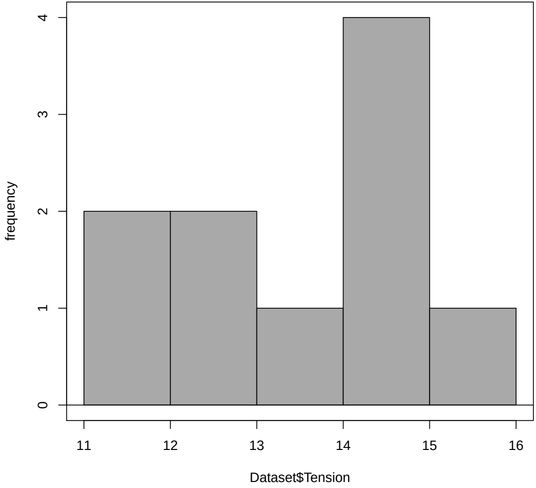
Exemple (suite)



Exemple (suite)



Exemple (suite)



Exemple (suite)

- On observe :

Sujet	1	2	3	4	5	6	7	8	9	10
x (ans)	56	42	72	36	63	47	55	49	38	42
y (mmHg)	14.7	12.5	16.0	11.8	14.9	12.8	15.0	14.5	11.5	14.0

$\rightarrow$ affectation des rangs

$\rightarrow$ calcul des différences de rangs

Sujet	1	2	3	4	5	6	7	8	9	10
r_X	8	3.5	10	1	9	5	7	6	2	3.5
r_Y	7	3	10	2	8	4	9	6	1	5
d	1	0.5	0	-1	1	1	-2	0	1	-1.5
d^2	1	0.25	0	1	1	1	4	0	1	2.25

Exemple (suite)

- Coefficient de corrélation estimé entre X et Y :

$$r = 1 - \frac{6 * 11.5}{10 * (10^2 - 1)} \approx 0.93$$

→ $|r| > 0.5 \Rightarrow$ Forte corrélation : Il semble que la tension artérielle systolique augmente avec l'âge chez les femmes

Peut-on conclure que le coefficient de corrélation linéaire ρ de la population est significativement différent de 0 ?

Exemple (suite)

- Coefficient de corrélation estimé entre X et Y :

$$r = 1 - \frac{6 * 11.5}{10 * (10^2 - 1)} \approx 0.93$$

$\rightarrow$ $|r| > 0.5 \Rightarrow$ Forte corrélation : Il semble que la tension artérielle systolique augmente avec l'âge chez les femmes

Peut-on conclure que le coefficient de corrélation linéaire ρ de la population est significativement différent de 0 ?

Test de corrélation (1)

étape 1 : Choix des hypothèses

question d'intérêt $\Rightarrow$ test bilatéral

$H_0 : \rho = 0$ $\Leftrightarrow$ on suppose l'indépendance entre X et Y
$H_1 : \rho \neq 0$

question d'intérêt $\Rightarrow$ test unilatéral

$H_0 : \rho = 0$ $\Leftrightarrow$ on suppose l'indépendance entre X et Y
$H_1 : \rho > 0$ (ou bien $<$)

Test de corrélation (2)

étape 2 : Choix du test, statistique de test et loi sous H_0

2 cas de figure :

- $n > 10$
- $n \leq 10$

Test de corrélation (3) - $n > 10$

étape 2 : Choix du test, statistique de test et loi sous H_0

Sous H_0 , avec hypothèse de **normalité** des 2 v.a., la statistique de test :

$$T = \frac{R - \rho}{\sqrt{Var(R)}} \sim Student_{(n-2)ddl}$$

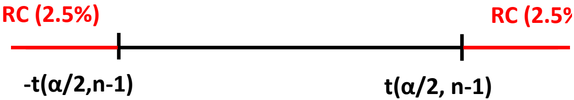
où $\sqrt{Var(R)} = \frac{1-R^2}{n-2}$

$$\Rightarrow T = \frac{R\sqrt{n-2}}{\sqrt{1-R^2}} \sim Student_{(n-2)ddl}$$

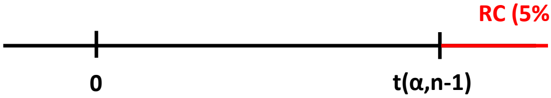
Test de corrélation (4) - $n > 10$

étape 3 : Choix du niveau de signification α, Définition de la région critique (RC)

α, test bilatéral, loi de Student



α, test unilatéral, loi de Student



Test de corrélation (5) - $n > 10$

étape 4 : Calcul de la statistique de test

Analyse numérique réalisée à l'aide des estimations issues des échantillons A et B

$$t_{exp} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

Test de corrélation (6) - $n > 10$

étape 5 & 6 : Conclusion statistique et épidémiologique

- Si $t_{exp} \in RC \rightarrow p_c < \alpha$.
  - Rejet de H_0 car moins de α % de chance qu'elle soit vraie.
  - Il semble que ρ soit $\neq$ de 0.
  - Les v.a. X et Y semblent **linéairement** corrélées.
- Si $t_{exp} \notin RC \rightarrow p_c > \alpha$.
  - Non rejet de H_0 car plus de α % de chance qu'elle soit vraie.
  - On ne peut pas montrer que ρ soit $\neq$ de 0.
  - On ne peut pas montrer que les v.a. X et Y semblent **linéairement** corrélées.

Remarque : p_c=probabilité de se tromper en rejetant H_0

Test de corrélation (7) - $n \leq 10$

- On range la v.a. X : (R_{Xi})
- Sous H_0 (hypothèse d'aucune corrélation),

  $\rightarrow$ le rang de la variable Y est supposé distribué aléatoire selon les $n!$ arrangements possibles
- Comparaison immédiate du coefficient de corrélation estimé r avec ceux théoriques calculés sous H_0 (fournies dans la table) à un certain seuil α

  Attention au nb de ddl= n-2 ! !

Table

α / n	0.50	0.20	0.10	0.05	0.02	0.01	0.005	0.002	0.001
4	0.600	1.000	1.000						
5	0.500	0.800	0.900	1.000	1.000				
6	0.371	0.657	0.829	0.886	0.943	1.000	1.000		
7	0.321	0.571	0.714	0.786	0.893	0.929	0.964	1.000	1.000
8	0.310	0.524	0.643	0.738	0.833	0.881	0.905	0.952	0.976
9	0.267	0.483	0.600	0.700	0.783	0.833	0.867	0.917	0.933
10	0.248	0.455	0.564	0.648	0.745	0.794	0.830	0.879	0.903
11	0.236	0.427	0.536	0.618	0.709	0.755	0.800	0.845	0.873
12	0.224	0.406	0.503	0.587	0.671	0.727	0.776	0.825	0.860
13	0.209	0.385	0.484	0.560	0.648	0.703	0.747	0.802	0.835
14	0.200	0.367	0.464	0.538	0.622	0.675	0.723	0.776	0.811
15	0.189	0.354	0.443	0.521	0.604	0.654	0.700	0.754	0.786
16	0.182	0.341	0.429	0.503	0.582	0.635	0.679	0.732	0.765
17	0.176	0.328	0.414	0.485	0.566	0.615	0.662	0.713	0.748
18	0.170	0.317	0.401	0.472	0.550	0.600	0.643	0.695	0.728
19	0.165	0.309	0.391	0.460	0.535	0.584	0.628	0.677	0.712
20	0.161	0.299	0.380	0.447	0.520	0.570	0.612	0.662	0.696
21	0.156	0.292	0.370	0.435	0.508	0.556	0.599	0.648	0.681
22	0.152	0.284	0.361	0.425	0.496	0.544	0.586	0.634	0.667
23	0.148	0.278	0.353	0.415	0.486	0.532	0.573	0.622	0.654
24	0.144	0.271	0.344	0.406	0.476	0.521	0.562	0.610	0.642
25	0.142	0.265	0.337	0.398	0.466	0.511	0.551	0.598	0.630
26	0.138	0.259	0.331	0.390	0.457	0.501	0.541	0.587	0.619
27	0.136	0.255	0.324	0.382	0.448	0.491	0.531	0.577	0.608
28	0.133	0.250	0.317	0.375	0.440	0.483	0.522	0.567	0.598
29	0.130	0.245	0.312	0.368	0.433	0.475	0.513	0.558	0.589
30	0.128	0.240	0.306	0.362	0.425	0.467	0.504	0.549	0.580
31	0.126	0.236	0.301	0.356	0.418	0.459	0.496	0.541	0.571
32	0.124	0.232	0.296	0.350	0.412	0.452	0.489	0.533	0.563
33	0.121	0.229	0.291	0.345	0.405	0.446	0.482	0.525	0.554
34	0.120	0.225	0.287	0.340	0.399	0.439	0.475	0.517	0.547
35	0.118	0.222	0.283	0.335	0.394	0.433	0.468	0.510	0.539
36	0.116	0.219	0.279	0.330	0.388	0.427	0.462	0.504	0.533
37	0.114	0.216	0.275	0.325	0.383	0.421	0.456	0.497	0.526
38	0.113	0.212	0.271	0.321	0.378	0.415	0.450	0.491	0.519
39	0.111	0.210	0.267	0.317	0.373	0.410	0.444	0.485	0.513
40	0.110	0.207	0.264	0.313	0.368	0.405	0.439	0.479	0.507

Test de corrélation (8) - $n \leq 10$

Conclusion statistique et épidémiologique

- Si $r > r_{theo} \rightarrow p_c < \alpha$.
  - Rejet de H_0 car moins de α % de chance qu'elle soit vraie.
  - Il semble que ρ soit $\neq$ de 0.
  - Les v.a. X et Y semblent **linéairement** corrélées.
- Si $r \leq r_{theo} \rightarrow p_c > \alpha$.
  - Non rejet de H_0 car plus de α % de chance qu'elle soit vraie.
  - On ne peut pas montrer que ρ soit $\neq$ de 0.
  - On ne peut pas montrer que les v.a. X et Y semblent **linéairement** corrélées.

Remarque : p_c=probabilité de se tromper en rejetant H_0

Exemple (suite)

étape 1 : Choix des hypothèses

question d'intérêt :

Exemple (suite)

étape 1 : Choix des hypothèses

question d'intérêt :

On souhaite tester si la tension artérielle systolique diffère avec l'âge chez les femmes

Exemple (suite)

étape 1 : Choix des hypothèses

question d'intérêt :

On souhaite tester si la tension artérielle systolique diffère avec l'âge chez les femmes

$\Rightarrow$ test bilatéral

$H_0 : \rho = 0 \quad \Leftrightarrow$ on suppose l'indépendance entre X et Y

$H_1 : \rho \neq 0$

Exemple (suite)

On a dans notre exemple : $n = 10$

$$r \approx 0.933$$

et pour $10 - 2$ ddl, avec $\alpha = 5\%$, on a :

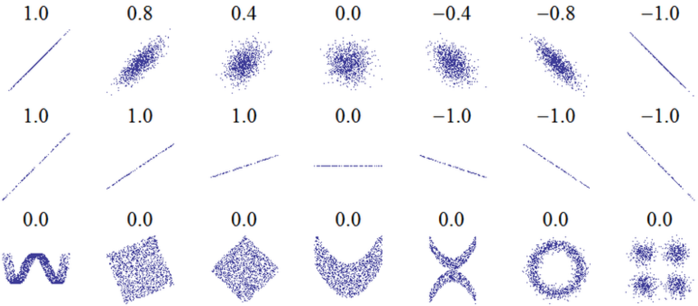
$$r_{theo} = 0.738$$

- $r > r_{theo} \rightarrow p_c < \alpha$
  - Rejet de H_0 car moins de α % de chance qu'elle soit vraie.
  - Il semble que ρ soit $\neq$ de 0.
  - Les v.a. X et Y semblent **linéairement** corrélées

$\Rightarrow$ la tension artérielle systolique semble liée à l'âge chez les femmes

http://www.r-project.org/

Quelques exemples/limites



- Attention aux conclusions non-valides pour une relation non-linéaire.
- Plusieurs relations possibles pour un même coefficient de corrélation → quantification de la relation souvent nécessaire.

Un peu de discussion...

- Coefficient de corrélation linéaire mesure la liaison entre 2 v.a. quantitatives

$\Longrightarrow$ **Attention** dans l'interprétion de la corrélation

- 1 corrélation entre 2 v.a. n'entraine pas obligatoirement une relation de causalité

$\longrightarrow$ 1 corrélation forte et significative peut signifier

- $\rightarrow$ un lien fortuit
- $\rightarrow$ un lien commun
- $\rightarrow$ un lien de causalité

Plan

Introduction
Comparaison de moyennes
2 échantillons indépendants
+ de 2 échantillons indépendants
2 échantillons appariés
Comparaison de fréquences
2 échantillons indépendants
2 échantillons appariés
Corrélation linéaire
χ^2 d'indépendance
Conclusion

Comparaison de 2 fréquences ou + :

Test du χ^2 d'indépendance

- On observe un échantillon de N sujets avec pour chaque sujet 2 critères de qualification :
  - 1 variable catégorielle A à k modalités
  - 1 variable catégorielle B à l modalités

$\Rightarrow$ Classification sous forme de tableau de contingence :

	B_1	...	B_j	...	B_l	Totaux lignes T_i
A_1	O_{11}	...	O_{1j}	...	O_{1l}	$O_{1.} = T_1$
...	...	...	...	...	...	...
A_i	O_{i1}	...	O_{ij}	...	O_{il}	$O_{i.} = T_i$
...	...	...	...	...	...	...
A_k	O_{k1}	...	O_{kj}	...	O_{kl}	$O_{k.} = T_k$
Totaux colonnes S_j	$O_{.1} = S_1$	...	$O_{.j} = S_j$	...	$O_{.l} = S_l$	$O_{..} = N$ =Effectif total

- Hypothèses :
  - H_0 : indépendance entre A et B
  - H_1 : dépendance entre A et B

- On peut montrer que la statistique de test suivante :

$$\chi^2_{exp} = \sum_i \sum_j \frac{(O_{ij} - C_{ij})^2}{C_{ij}} \sim \chi^2_{(k-1)(l-1)} \ ddl$$

où C_{ij} : effectif théorique sous H_0

- Conditions de validité :

$C_{ij} \geq 5$ sinon regrouper (*éventuellement*) des classes adjacentes

Région critique :

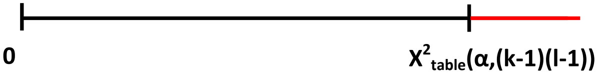


TABLE DE LA FONCTION DE REPARTITION DU χ^2

Degré liberté	Probabilité α 0,995	0,990	0,975	0,950	0,900	0,100	0,050	0,025	0,010	0,005
1	0,00	0,00	0,00	0,00	0,02	2,71	3,84	5,02	6,63	7,88
2	0,01	0,02	0,05	0,10	0,21	4,61	5,99	7,38	9,21	10,60
3	0,07	0,11	0,22	0,35	0,58	6,25	7,81	9,35	11,34	12,84
4	0,21	0,30	0,48	0,71	1,06	7,78	9,49	11,14	13,28	14,86
5	0,41	0,55	0,83	1,15	1,61	9,24	11,07	12,83	15,09	16,75
6	0,68	0,87	1,24	1,64	2,20	10,64	12,59	14,45	16,81	18,55
7	0,99	1,24	1,69	2,17	2,83	12,02	14,07	16,01	18,48	20,28
8	1,34	1,65	2,18	2,73	3,49	13,36	15,51	17,53	20,09	21,95
9	1,73	2,09	2,70	3,33	4,17	14,68	16,92	19,02	21,67	23,59
10	2,16	2,56	3,25	3,94	4,87	15,99	18,31	20,48	23,21	25,19
11	2,60	3,05	3,82	4,57	5,58	17,28	19,68	21,92	24,73	26,76
12	3,07	3,57	4,40	5,23	6,30	18,55	21,03	23,34	26,22	28,30
13	3,57	4,11	5,01	5,89	7,04	19,81	22,36	24,74	27,69	29,82
14	4,07	4,66	5,63	6,57	7,79	21,06	23,68	26,12	29,14	31,32
15	4,60	5,23	6,26	7,26	8,55	22,31	25,00	27,49	30,58	32,80
16	5,14	5,81	6,91	7,96	9,31	23,54	26,30	28,85	32,00	34,27
17	5,70	6,41	7,56	8,67	10,09	24,77	27,59	30,19	33,41	35,72
18	6,26	7,01	8,23	9,39	10,86	25,99	28,87	31,53	34,81	37,16
19	6,84	7,63	8,91	10,12	11,65	27,20	30,14	32,85	36,19	38,58

Conclusion :

- $\chi^2_{exp} \geq \chi^2_{table(\alpha,(k-1)l-1)\ ddl} \rightarrow p < \alpha$

$\Rightarrow$ Rejet de H_0 : Dépendance entre les variables A et B

- $\chi^2_{exp} < \chi^2_{table(\alpha,(k-1)l-1)\ ddl} \rightarrow p > \alpha$

$\Rightarrow$ Non rejet de H_0 : Nous n'avons pas pu mettre en évidence une dépendance entre les variables A et B

Comparaison de 2 fréquences ou + :
Exemple

- Lors d'un essai clinique comparant 2 traitements A et B (150 patients randomisés dans chaque bras), l'ensemble des effets indésirables sont collectés
- Les patients sont classés en 3 groupes : pas d'effet indésirable (PEI), au moins un effet indésirable (et aucun grave) (EI), au moins un effet indésirable grave (EIG)
- Voici la répartition des sujets en fonction de leur traitement et de la survenue d'effets indésirables :

	PEI	EI	EIG
traitement A	71	50	29
traitement B	98	32	20

Comparaison de 2 fréquences ou + :

Exemple

On cherche à répondre à la question suivante :
"Existe-t-il un lien significatif entre le traitement et les effets secondaires ?" (risque d'erreur de 1ère espèce fixé à 5%) ?

- Quelle est l'hypothèse nulle H_0 de ce test ?
- Quelle est l'hypothèse alternative H_1 de ce test ?

Comparaison de 2 fréquences ou + :
Exemple

On cherche à répondre à la question suivante :
"Existe-t-il un lien significatif entre le traitement et les effets secondaires ?" (risque d'erreur de 1ère espèce fixé à 5%) ?

- Quelle est l'hypothèse nulle H_0 de ce test ?
- Quelle est l'hypothèse alternative H_1 de ce test ?

H_0 : les proportions d'effets secondaires sont égales dans les 2 groupes de traitement

H_1 : Il y a au moins 2 proportions qui sont différentes

Comparaison de 2 fréquences ou + :

Exemple

- Quel type de test permet de répondre à cette question ?

Comparaison de 2 fréquences ou + :

Exemple

- Quel type de test permet de répondre à cette question ?

Test de χ^2 d'indépendance

$\rightarrow$ Test de l'existence d'un lien entre le traitement et la survenue d'effets secondaires

$\rightarrow$ c'est-à-dire l'indépendance entre la survenue d'effets secondaires et la prise du traitement A ou B

Comparaison de 2 fréquences ou + :
Exemple

- Calcul de la statistique de test

Tableau de contingence des effectifs théoriques :

	PEI	EI	EIG	
traitement A	$\frac{150*169}{300} = 84.5$	41	24.5	150
traitement B	84.5	41	24.5	150
Total	169	82	49	300

Sous H_0,

- la proportion d'effets indésirables graves(...) devrait être la même quelque soit le traitement :

$$49/300 \approx 0.1633$$

- rapportée aux patients traités par le médicament A :

$$150 * 49/300 = 24.5$$

Comparaison de 2 fréquences ou + :

Exemple

Conditions d'application :

$$\frac{49 * 150}{300} = 24.5 > 5$$

La statistique de test est égale à

$$\chi^2_{calc} = \sum_{i,j} \frac{(O_{ij} - C_{ij})^2}{C_{ij}} \sim \chi^2_{theo}(2ddl)$$

Tableau des écarts au carré $\frac{(O_{ij} - C_{ij})^2}{C_{ij}}$:

	PEI	EI	EIG
traitement A	2.16	1.98	0.83
traitement B	2.16	1.98	0.83

La statistique de test calculée est :

$$\chi^2_{calc} = 2.16 + 2.16 + 1.98 + 1.98 + 0.83 + 0.83 = 9.94$$

- Conclusion

La région critique au seuil $\alpha = 0.05$ est définit en fonction du nombre de degré de liberté

$$RC = [5.99; +\infty[$$

$$\chi^2_{calc} = 9.94 \in RC$$

$\Rightarrow$ Rejet de H_0 au seuil $\alpha = 0.05$

Il semble donc que les proportions des différents types d'effets indésirables ne soient pas indépendantes de la prescription médicale ($p_c < 0.05$).

Comparaison de 2 fréquences ou + :

Exemple sous R

http://www.r-project.org/

Comparaison de 2 fréquences ou + :

Exemple sous R


```
> t<-table(baseChisq$Traitement,baseChisq$EffInd)
> t

    EI EIG PEI
  A 50  29  71
  B 32  20  98
>
> test<-chisq.test(t)
> test

        Pearson's Chi-squared test

data:  t
X-squared = 9.9179, df = 2, p-value = 0.00702

>
> test$expected

    EI  EIG  PEI
  A 41 24.5 84.5
  B 41 24.5 84.5
```

Comparaison de 2 fréquences ou + :

Test du χ^2 d'indépendance

- Cas particulier d'un tableau 2*2 :

	Malades	Non Malade	Total
Exposés	a	b	a+b
Non exposés	c	d	c+d
Total	a+c	b+d	N

$\Rightarrow$ Comparaison de 2 % sur 2 grands échantillons indépendants

- H_0 : le % de bronchite est le même quelle que soit l'exposition

$$H_0 : \pi_e = \pi_{\bar{e}} \quad \text{et} \quad H_1 : \pi_e \neq \pi_{\bar{e}}$$

où π_e : % de bronchite chez les exposés

où $\pi_{\bar{e}}$: % de bronchite chez les non exposés

Notion d'Odds Ratio

- Mesure d'association entre un facteur d'exposition (facteur de risque) et la présence de la maladie

	Malades (M)	Non Malade ($\bar{M}$)	Total
Exposés (E)	a	b	a+b
Non exposés ($\bar{E}$)	c	d	c+d
Total	a+c	b+d	N

Notion d'Odds Ratio

- Odds (côte) :

$$\frac{Prob(M)}{Prob(\bar{M})} = \frac{p}{1-p}$$

- Odds chez les exposés :

$$\frac{a/(a+b)}{b/(a+b)} = \frac{a}{b}$$

- Odds chez les non exposés :

$$\frac{c/(c+d)}{d/(c+d)} = \frac{c}{d}$$

$\Rightarrow$ Odds ratio :

$$\frac{\text{Odds chez les exposés}}{\text{Odds chez les non exposés}} = \frac{a/b}{c/d} = \frac{ad}{bc}$$

Notion d'Odds Ratio

Remarques :

- Si Odds Ratio=OR>1 $\Rightarrow$ facteur de risque
- Si Odds Ratio=OR<1 $\Rightarrow$ facteur protecteur

$\Rightarrow$ OR est-il significativement $\neq$ de 1 ?

Intervalle de confiance, test

Odds Ratio

Exemple

- On observe le tableau de contingence suivant chez 460 patients chez lesquels on s'intéresse à l'efficacité d'un vaccin contre la grippe

	Grippe	Pas de grippe	Total
Vaccin	20	220	240
Placebo	80	140	220
Total	100	360	460

$\Rightarrow$ Existe-t-il un lien entre la grippe et la vaccination ? (au risque de 1ère espèce $\alpha = 5\%$?

$\Rightarrow$ Test du χ^2

- Calcul de la statistique de test

Tableau de contingence des effectifs théoriques :

	Grippe	Pas de grippe	Total
Vaccin	$\frac{240*100}{460} = 52.17$	187.83	240
Placebo	47.83	172.17	220
Total	100	360	460

Sous H_0,

- la proportion de grippe devrait être la même quelque soit le genre :

$$100/460 \approx 0.22$$

- rapportée au nombre de personnes vaccinées :

$$240 * 100/460 = 52.17$$

Odds Ratio

Exemple

Conditions d'application :

$$\frac{220 * 100}{460} = 47.83 > 5$$

La statistique de test est égale à

$$\chi^2_{calc} = \sum_{i,j} \frac{(O_{ij} - C_{ij})^2}{C_{ij}} \sim \chi^2_{theo}(2ddl)$$

Tableau des écarts au carré $\frac{(O_{ij} - C_{ij})^2}{C_{ij}}$:

	Grippe	Pas de grippe
Vaccin	19.84	5.51
Placebo	21.64	6.01

La statistique de test calculée est :

$$\chi^2_{calc} = 19.84 + 5.51 + 21.64 + 6.01 \approx 53$$

Conclusion

La région critique au seuil $\alpha = 0.05$ est définit en fonction du nombre de degré de liberté

$$RC = [3.84; +\infty[$$

$$\chi^2_{calc} = 53 \in RC$$

$\Rightarrow$ Rejet de H_0 car moins de α % de chance qu'elle soit vraie.

Il semble que la proportion de grippe soit différente selon que l'on soit vacciné ou non ($p_c < 0.05$).

$\Rightarrow$ Existence d'un lien entre grippe et vaccination

Conclusion

La région critique au seuil $\alpha = 0.05$ est définit en fonction du nombre de degré de liberté

$$RC = [3.84; +\infty[$$

$$\chi^2_{calc} = 53 \in RC$$

$\Rightarrow$ Rejet de H_0 car moins de α % de chance qu'elle soit vraie.

Il semble que la proportion de grippe soit différente selon que l'on soit vacciné ou non ($p_c < 0.05$).

$\Rightarrow$ Existence d'un lien entre grippe et vaccination

$\Rightarrow$ Intensité de ce lien ?

Odds Ratio

$$OR = \frac{20 * 140}{220 * 80} = 0.159 < 1$$

$\Rightarrow$ La vaccination apparaît comme un facteur protecteur de la grippe

Intervalle de confiance d'un OR

- La distribution de l'OR n'est pas symétrique ($OR > 0$)
  $\rightarrow$ Distribution non normale

$\Rightarrow$ $\ln(OR) \sim$ distribution normale (approximativement)

- Calcul de l'IC de $\ln(OR)$, puis on en déduit celui de OR

- Variance de $\ln(OR)$:

$$s^2(\ln(OR)) = \frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}$$

$\Rightarrow$ Intervalle de confiance à 95% de $\ln(OR)$:

$$IC_{95\%}(\ln(OR)) = \ln(OR) \pm 1.96 * s(\ln(OR))$$

$\Rightarrow$ Intervalle de confiance à 95% de OR :

$$IC_{95\%}(OR) = e^{\ln(OR) \pm 1.96 * s(\ln(OR))}$$

Intervalle de confiance d'un OR

Exemple

- $OR = 0.159 \rightarrow \ln(OR) = \ln(0.159) \approx -1.839$
- $s^2(\ln(OR)) = \frac{1}{20} + \frac{1}{220} + \frac{1}{80} + \frac{1}{140} \approx 0.0742$
- dans l'exemple :

$$IC_{95\%}(OR) = e^{-1.839 \pm 1.96 * 0.0742} = [0.093; 0.272]$$

- Existe-il un lien entre grippe et vaccination ?

Intervalle de confiance d'un OR

Exemple

- $OR = 0.159 \rightarrow \ln(OR) = \ln(0.159) \approx -1.839$
- $s^2(\ln(OR)) = \frac{1}{20} + \frac{1}{220} + \frac{1}{80} + \frac{1}{140} \approx 0.0742$
- dans l'exemple :

$$IC_{95\%}(OR) = e^{-1.839 \pm 1.96 * 0.0742} = [0.093; 0.272]$$

- Existe-il un lien entre grippe et vaccination ?

$$1 \notin IC_{95\%}(OR)$$

$$\Rightarrow \text{Rejet de } H_0 \text{ au seuil 5\% (supposant } OR = 1)$$

Risque Relatif et Odds Ratio

- Risque Relatif=RR : Rapport des risques

	Malades (M)	Non Malade ($\bar{M}$)	Total
Exposés (E)	a	b	a+b
Non exposés ($\bar{E}$)	c	d	c+d
Total	a+c	b+d	N

$$RR = \frac{Prob(M|E)}{Prob(M|\bar{E})} = \frac{a/(a+b)}{c/(c+d)}$$

$\Rightarrow$ RR s'interprète comme un coefficient de proportionnalité

par ex., $RR = 2 \Rightarrow$ les sujets exposés ont 2 fois plus de risque d'être malade que les non exposés

$\Rightarrow$ Interprétation extrêmement intéressante et pertinente

Risque Relatif et Odds Ratio

- Problème :

Risque Relatif et Odds Ratio

- Problème :
  - Nécessité d'avoir un échantillon représentatif

$\rightarrow$ Prévalence de la maladie doit être celle de la population

$$p = \frac{a+c}{N}$$

$\rightarrow$ pas sytématique

par ex. : étude cas-témoin où on a souvent

$$p = \frac{a+c}{N} = \frac{1}{2} \text{ car } (a+c) = (b+d)$$

$\Rightarrow$ Odds Ratio calculable ! !

Risque Relatif et Odds Ratio

- Problème :
  - Nécessité d'avoir un échantillon représentatif

  $\rightarrow$ Prévalence de la maladie doit être celle de la population

$$p = \frac{a+c}{N}$$

  $\rightarrow$ pas sytématique

  par ex. : étude cas-témoin où on a souvent

$$p = \frac{a+c}{N} = \frac{1}{2} \text{ car } (a+c) = (b+d)$$

$\Rightarrow$ Odds Ratio calculable ! !

- Si maladie rare, alors $OR \approx RR$

$$a << b \text{ et } c << d \Rightarrow RR \approx \frac{a/b}{c/d} = \frac{ad}{bc}$$

Plan

Conclusion

Un peu de discussion...

- p_c : probabilité critique ou degré de signification

$\Rightarrow$ probabilité de rejeter H_0 à tort

$\Rightarrow$ mesure la crédibilité de H_0, quantifie le “désacccord” entre ce qu'on observe et l'hypothèse nulle H_0

Mais p_c ne quantifie pas l'importance de la différence

Un peu de discussion...

Ne pas mettre en évidence de différence statistiquement significative entre 2 groupes

ne signifie pas

qu'il y ait **équivalence** entre les 2 groupes (H_0 vraie) (Altman, 2005)

$\Rightarrow$ Un résultat non statistiquement significatif peut avoir 2 causes :

- l'hypothèse H_0 est vraie (i.e. les 2 groupes proviennent de la même population
- la puissance statistique n'est pas suffisante (i.e. nombre de sujets insuffisant, β existe et $\neq 0$)

MERCI DE VOTRE ATTENTION