Chapitre 10

Alternatives non paramétriques aux tests statistiques sur deux moyennes

Tests paramétriques et non paramétriques

Tests paramétriques:

- Estimation de paramètres. Ex: μ, σ^2
- Condition d'application: normalité

Tests non paramétriques:

- Pas de condition d'application
- Moins puissants

Tester la condition de normalité

<u>Tests statistiques</u>:

- **Test W de Shapiro et Wilk (1965)**
- **Test K^2 de D'Agostino et Pearson (1971)**

<u>Description des données</u>:

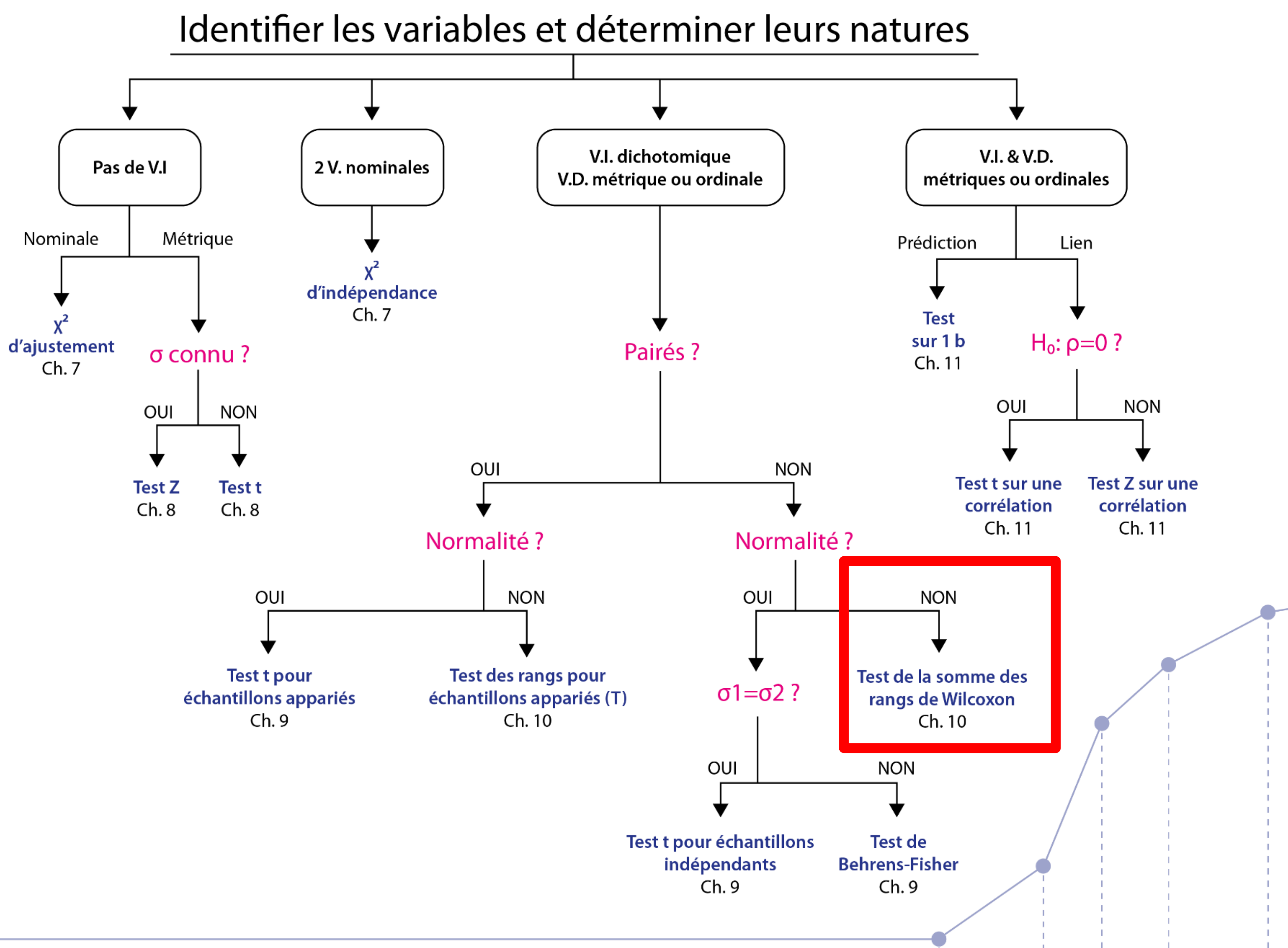
- **Tendance centrale ?**
- **Forme bimodale ?**
- **Effet plafond ou plancher**

Le test de la somme des rangs de Wilcoxon

Equivalent non paramétrique du test t sur deux moyennes indépendantes

Transformer les données en rangs

Choisir le bon test d'hypothèse



Exemple

Les patients cardiaques vivent-ils plus d’événements stressants dans les années qui précèdent un malaise cardiaque.

Deux groupes: patients cardiaques et contrôles

Exemple

Patients cardiaques :	32	8	7	29	5	0
Patients contrôles :	1	2	2	3	6	1

Exemple

H_0 : Les deux échantillons proviennent d'une population identique

H_A : Les deux échantillons proviennent de populations différentes

Exemple

Transformer les données en rangs :

32 8 7 29 5 0 1 2 2 3 6 1

0 1 1 2 2 3 5 6 7 8 29 32

1 2,5 2,5 4,5 4,5 6 7 8 9 10 11 12

Exemple

<u>Pour les ex-aequo</u> :

Attribuer le rang moyen

0	1	1	2	2	3	5	6	7	8	29	32
1	2,5	2,5	4,5	4,5	6	7	8	9	10	11	12

Ex : Rang moyen = (2 + 3)/2 = 2,5

	Patients cardiaques						Patients contrôles					
Données brutes	32	8	7	29	5	0	1	2	2	3	6	1
Rangs	12	10	9	11	7	1	2,5	4,5	4,5	6	8	2,5
$\sum$ rangs	50						28					

Exemple

<u>**Vérification des calculs**</u> **:**

1. Rang le plus élevé = N (sauf si ex-aequo)

2. Somme totale des rangs = N(N+1)/2

$$N(N+1)/2 = 12 \times 13 /2 = 78 = 50 + 28$$

Logique du test de Wilcoxon

<u>Si deux groupes équivalents</u>:

Sommes des rangs similaires

<u>Si deux groupes différents</u>:

Un groupe contient une somme importante et l'autre une somme très petite

Effectifs égaux / inégaux

A distinguer:

Effectifs égaux

$(n_1 = n_2)$

Effectifs inégaux

$(n_1 \neq n_2)$

Test de la somme des rangs de Wilcoxon pour effectifs égaux

Calculer W_s

<u>Test bilatéral</u>

W_S = la plus petite des deux sommes de rangs

<u>Test unilatéral</u>

W_S = somme des rangs qui serait la plus petite selon H_A

Calculer W_s

H_0 : Patients contrôles = patients cardiaques

H_A : Patients contrôles < patients cardiaques

$W_S = 28$

La table W_s

La table W_s contient les plus petites valeurs qu'on s'attend à trouver avec leur probabilité indiquée en haut du tableau

N_2	$N_1 = 6$					
	.001	.005	.010	.025	.05	.10
6		23	24	26	28	30
7	21	24	25	27	29	32
8	22	25	27	29	31	34
9	23	26	28	31	33	36
10	24	27	29	32	35	38
11	25	28	30	34	37	40
12	25	30	32	35	38	42
13	26	31	33	37	40	44

	$N_1 = 6$					
N_2	.001	.005	.010	.025	.05	.10
6		23	24	26	28	30
7	21	24	25	27	29	32
8	22	25	27	29	31	34
9	23	26	28	31	33	36
10	24	27	29	32	35	38
11	25	28	30	34	37	40
12	25	30	32	35	38	42
13	26	31	33	37	40	44

Exemple

H_0 : Patients contrôles = patients cardiaques

H_A : Patients contrôles < patients cardiaques

$W_S = 28$ $W_{s\ 0,05} = 28$

Rejeter H_0 si W_S <u>inférieure ou égale</u> à valeur critique

Exemple

$W_S = 28 \leq W_{s\ 0,05} = 28$

Rejeter H_0

Les patients cardiaques subissent plus d'événements stressants

Test de la somme des rangs de Wilcoxon pour effectifs inégaux

Les statistiques W_s et W_s'

W_s = La plus petite des deux sommes

Problème: En cas d'inégalité des N, la somme dépend aussi du N de chaque groupe

Les statistiques W_s et W_s'

Cardiaques (N = 2)			Contrôles (N = 5)				
	48	55	5	7	8	9	10
R	6	7	1	2	3	4	5

$\sum$ rangs = 13

$\sum$ rangs = 15

$W_s = 13$

Les statistiques W_s et W_s'

	Cardiaques (N = 2)		Contrôles (N = 5)				
	48	55	5	7	8	9	10
R	2	1	7	6	5	4	3

$\sum$ rangs = 3

$\sum$ rangs = 25

$W_s = 3$

Les statistiques W_s et W_s'

W_s = Somme des rangs du plus petit groupe (N le plus petit)

$W_s' = n_1(n_1 + n_2 + 1) - W_s$

$W_s' = 2\,(2 + 5 + 1) - 13 = 16 - 13 = 3$

Les statistiques W_s et W_s'

W_s = Somme des rangs du plus petit groupe (N le plus petit)

$W_s' = n_1(n_1 + n_2 + 1) – W_s$

<u>Attention : n_1 toujours le plus petit</u>

Faut-il utiliser W_s ou W_s'

Test bilatéral

Prendre la plus petite valeur de W_s ou W_s' et comparer à la valeur critique de la table

Attention: diviser α par 2

Faut-il utiliser W_s ou W_s'

Test unilatéral

Statistique à sélectionner dépend de H_A

- **Si échantillon avec n inférieur est supposé plus petit: prendre W_s**
- **Si échantillon avec n inférieur est supposé plus grand: prendre W_s'**

Faut-il utiliser W_s ou W_s'

Rejeter H_0 si la statistique sélectionnée (W_s ou W_s') est <u>inférieure ou égale</u> à valeur critique

Exemple

H_0 : Patients contrôles = patients cardiaques

H_A : Patients contrôles ≠ patients cardiaques

	Patients cardiaques						Patients contrôles				
Données brutes	32	8	7	29	5	0	1	2	2	3	6
Rangs	11	9	8	10	6	1	2	3,5	3,5	5	7
$\sum$ rangs	45						21				

Exemple

$W_s = 21$

$W_s' = n_1(n_1 + n_2 + 1) - W_s$

$W_s' = 5\,(6 + 5 + 1) - 21 = 39$

Test bilatéral : sélectionner la plus petite :

$W_s = 21$

N_2	$N_1 = 5$					
	.001	.005	.010	.025	.05	.10
5		15	16	17	19	20
6		16	17	18	20	22
7		16	18	20	21	23
8	15	17	19	21	23	25
9	16	18	20	22	24	27
10	16	19	21	23	26	28
11	17	20	22	24	27	30
12	17	21	23	26	28	32

	$N_1 = 5$					
N_2	.001	.005	.010	.025	.05	.10
5		15	16	17	19	20
6		16	17	18	20	22
7		16	18	20	21	23
8	15	17	19	21	23	25
9	16	18	20	22	24	27
10	16	19	21	23	26	28
11	17	20	22	24	27	30
12	17	21	23	26	28	32

Exemple

Puisque $W_s = 21 > 18$, ne pas rejeter H_0

Pas de preuves suffisantes pour affirmer que les patients cardiaques subissent plus d'événements stressants

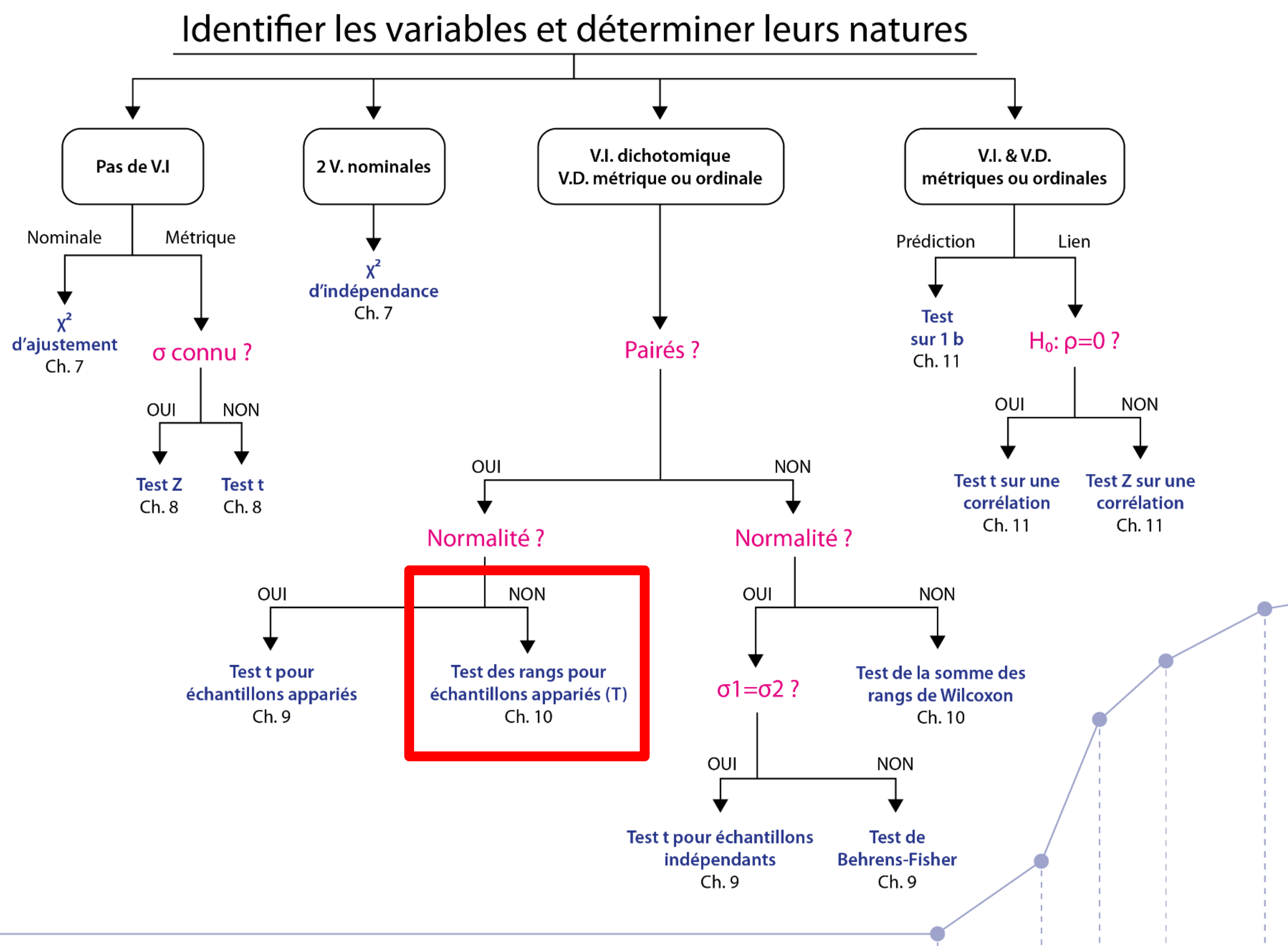
Test des rangs pour échantillons pairés de Wilcoxon

Le test des rangs pour échantillons pairés de Wilcoxon

Equivalent non paramétrique du test t pour échantillons pairés

Transformer les scores de différence en rangs

Choisir le bon test d'hypothèse



Exemple

Un entraînement sportif de six mois permet-il de modifier la tension artérielle ?

8 sujets: tension artérielle avant et après entraînement sportif

Exemple

H_0 : La tension artérielle est la même avant et après l’entraînement

H_A : La tension artérielle est différente avant et après l’entraînement

Avant	130	170	125	170	130	130	145	160
Après	120	163	120	135	143	136	144	120
Différence	10	7	5	35	-13	-6	1	40
Rangs D	5	4	2	7	6	3	1	8
Rangs signe	5	4	2	7	-6	-3	1	8

$$T_+ = \sum(\text{rangs positifs}) = 27$$

$$T_- = \sum(\text{rangs négatifs}) = 9$$

Remarques

1. <u>Scores de différence égaux</u>

   Donner des rangs ex-æquo

2. <u>Scores de différence = zéro</u>

   Ignorer ces données et réduire N

Logique du test

Si les deux ensembles de données sont identiques:

Somme rangs positifs

=

Somme rangs négatifs

La statistique T

Test bilatéral

T = La plus petite des deux sommes de rangs

Test unilatéral

T = Somme des rangs qui devait être la plus petite selon H_A

Exemple

$T_+ = \sum(\text{rangs positifs}) = 27$

$T_- = \sum(\text{rangs négatifs}) = 9$

$T_{obs} = 9$ (test bilatéral)

n	α :	0.25	0.10	0.05	0.025	0.01
5		4	2	0		
6		6	3	2	0	
7		9	5	3	2	0
8		12	8	5	3	1
9		16	10	8	5	3
10		20	14	10	8	5
11		24	17	13	10	7
…	…	…	…	…	…	…

n	α :	0.25	0.10	0.05	0.025	0.01
5		4	2	0		
6		6	3	2	0	
7		9	5	3	2	0
8		12	8	5	3	1
9		16	10	8	5	3
10		20	14	10	8	5
11		24	17	13	10	7
…	…	…	…	…	…	…

Exemple

T_{obs} = 9 (test bilatéral)

n = 8

α = 0,05

$T_{0,025}$ = 3

Rejeter H_0 si $T_{obs} \leq T_{0,025}$

Exemple

Puisque 9 > 3, ne pas rejeter H_0.

Nous n'avons pas suffisamment de preuves pour affirmer qu'un entraînement sportif de 6 mois réduit la tension artérielle.

Choisir le bon test d'hypothèse