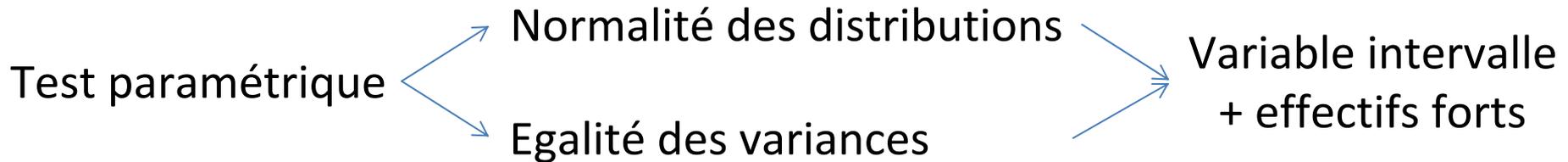


Tests non paramétriques



Nonparametric, or distribution free tests are so-called because the assumptions underlying their use are “fewer and weaker than those associated with parametric tests” (*Siegel & Castellan, 1988, p. 34*).

Un test non paramétrique est un test dont le modèle ne précise pas les conditions que doivent remplir les paramètres de la population dont a été extrait l'échantillon (*Ramousse, Le Berre & Le Guelte, 1996*).

Avantages/Inconvénients des Tests non paramétriques

Avantages

1. Leur emploi se justifie lorsque les conditions d'applications des autres méthodes ne sont pas satisfaites, même après d'éventuelles transformation de variables.
2. Pour des échantillons de taille très faible (jusqu'à $n = 6$ voire 4), la seule possibilité est l'utilisation d'un test non paramétrique (sauf si la nature exacte de la distribution de la population est précisément connue). Ceci permet une diminution du coût ou du temps nécessaire à la collecte des informations.
3. Il existe des tests non paramétriques permettant de traiter des échantillons composés à partir d'observations provenant de populations différentes. De telles données ne peuvent être traitées par les tests paramétriques sans faire des hypothèses irréalistes.
4. Seuls des tests non paramétriques permettent le traitement de données qualitatives : soit exprimées en rangs ou en plus ou moins (échelle ordinale), soit nominales.

Inconvénients

1. Les tests paramétriques, quand leurs conditions sont remplies, sont plus puissants que les tests non paramétriques.

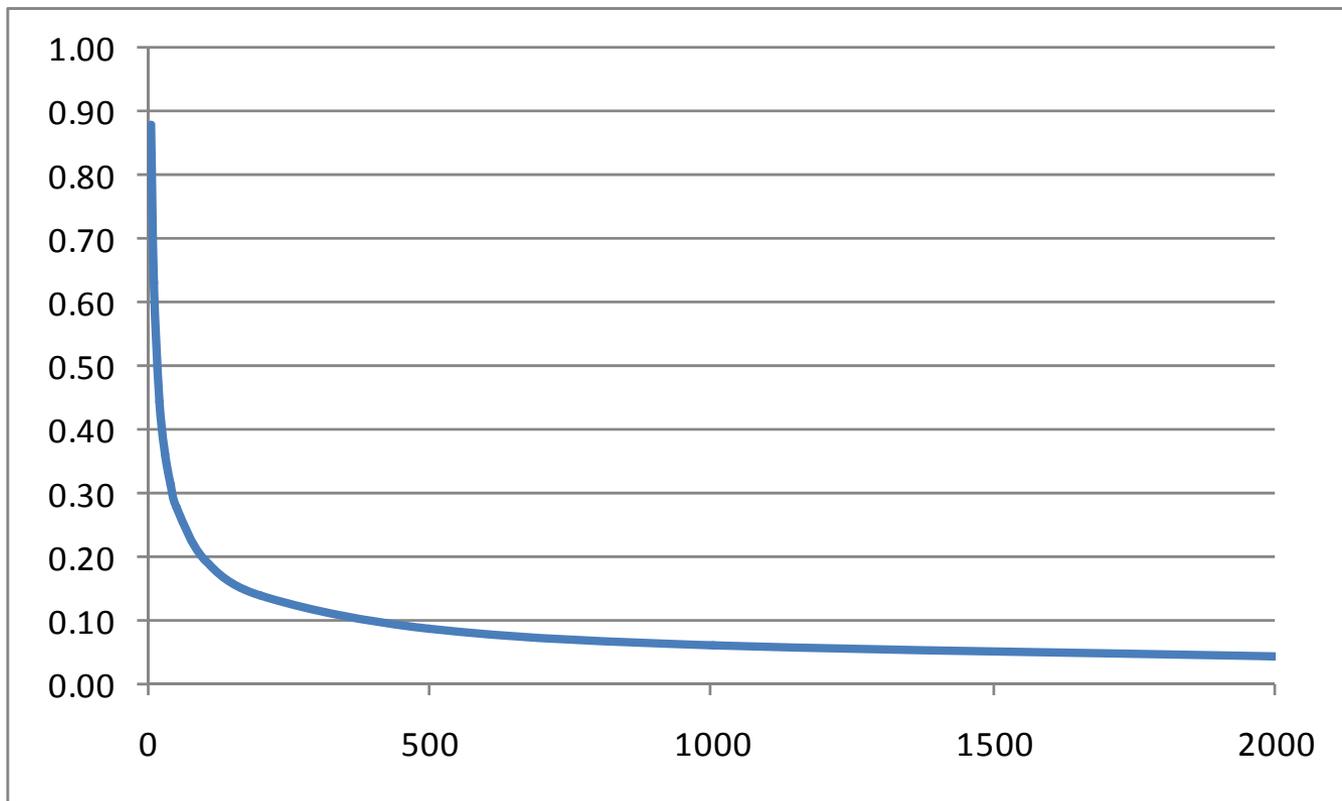
(Ramousse, Le Berre & Le Guelte, 1996)

Quelques généralités sur les Tests statistiques

$$t = \frac{\sqrt{n-2} \cdot r}{\sqrt{1-r^2}}$$

$$r = \pm \frac{t}{\sqrt{t^2 + n - 2}}$$

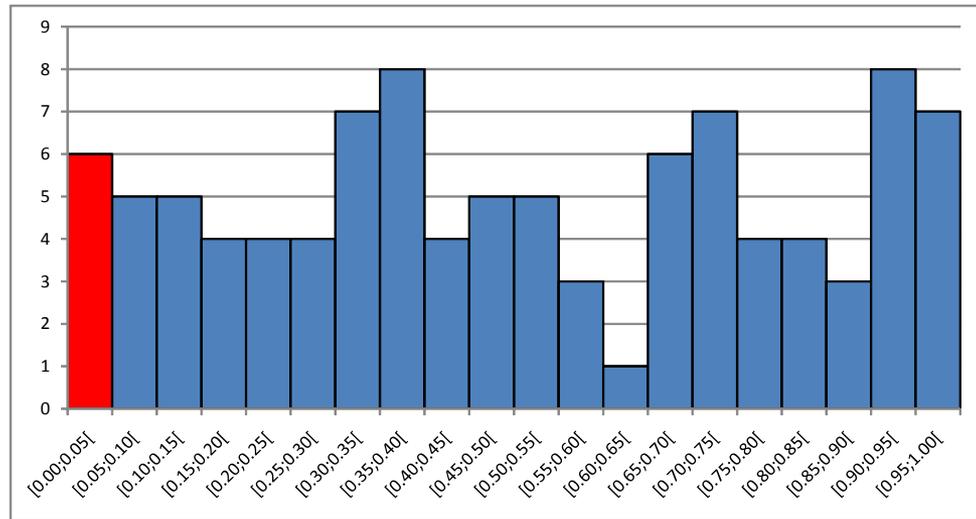
n	5	10	20	30	40	50	100	200	500	1000	2000
r	0.878	0.632	0.444	0.361	0.312	0.279	0.197	0.139	0.088	0.062	0.044
t	3.182	2.306	2.101	2.048	2.024	2.011	1.984	1.972	1.965	1.962	1.961



Quelques généralités sur les Tests statistiques

Illustration Test du Chi²

<i>i</i>	Groupe A	Groupe B												
			10.25%											
1	Groupe A_2	Groupe B_3	1	21.30%	Groupe A_1	5	8	11	12	36				
2	Groupe A_2	Groupe B_2	2	35.45%	Groupe A_2	16	17	9	13	55	0.05	[0.00;0.05[6	
3	Groupe A_3	Groupe B_2	3	72.49%	Groupe A_3	11	19	12	12	54	0.10	[0.05;0.10[5	
4	Groupe A_2	Groupe B_3	4	28.44%	Groupe A_4	21	15	13	6	55	0.15	[0.10;0.15[5	
5	Groupe A_2	Groupe B_2	5	71.77%	Total	53	59	45	43	200	0.20	[0.15;0.20[4	
6	Groupe A_1	Groupe B_3	6	12.23%							0.25	[0.20;0.25[4	
7	Groupe A_1	Groupe B_3	7	2.02%							0.30	[0.25;0.30[4	
8	Groupe A_4	Groupe B_3	8	82.44%	Groupe A_1	9.54	10.62	8.10	7.74	36	0.35	[0.30;0.35[7	
9	Groupe A_2	Groupe B_3	9	19.19%	Groupe A_2	14.58	16.23	12.38	11.83	55	0.40	[0.35;0.40[8	
10	Groupe A_4	Groupe B_3	10	12.60%	Groupe A_3	14.31	15.93	12.15	11.61	54	0.45	[0.40;0.45[4	
11	Groupe A_1	Groupe B_1	11	88.24%	Groupe A_4	14.58	16.23	12.38	11.83	55	0.50	[0.45;0.50[5	
12	Groupe A_2	Groupe B_4	12	34.85%	Total	53	59	45	43	200	0.55	[0.50;0.55[5	
13	Groupe A_4	Groupe B_2	13	80.62%							0.60	[0.55;0.60[3	
14	Groupe A_4	Groupe B_1	14	50.65%							0.65	[0.60;0.65[1	
15	Groupe A_3	Groupe B_1	15	66.47%	Groupe A_1	2.16	0.65	1.04	2.34		0.70	[0.65;0.70[6	
16	Groupe A_4	Groupe B_2	16	96.70%	Groupe A_2	0.14	0.04	0.92	0.12		0.75	[0.70;0.75[7	
17	Groupe A_3	Groupe B_2	17	79.93%	Groupe A_3	0.77	0.59	0.00	0.01		0.80	[0.75;0.80[4	
18	Groupe A_1	Groupe B_3	18	45.87%	Groupe A_4	2.83	0.09	0.03	2.87		0.85	[0.80;0.85[4	
19	Groupe A_4	Groupe B_1	19	72.75%							0.90	[0.85;0.90[3	
20	Groupe A_1	Groupe B_2	20	46.01%							0.95	[0.90;0.95[8	
21	Groupe A_4	Groupe B_3	21	97.17%	Fractile	16.92					1.00	[0.95;1.00[7	
22	Groupe A_4	Groupe B_3	22	56.47%	Proba. Crit.	10.25%								
23	Groupe A_1	Groupe B_4	23	69.45%									Total	100
24	Groupe A_2	Groupe B_2	24	73.86%										
25	Groupe A_3	Groupe B_4	25	62.23%										
26	Groupe A_1	Groupe B_3	26	44.57%										
27	Groupe A_3	Groupe B_4	27	34.73%										
28	Groupe A_2	Groupe B_1	28	79.44%										
29	Groupe A_2	Groupe B_2	29	2.24%										
30	Groupe A_4	Groupe B_3	30	79.23%										
31	Groupe A_4	Groupe B_2	31	35.16%										
32	Groupe A_4	Groupe B_4	32	66.69%										
33	Groupe A_1	Groupe B_4	33	19.69%										
34	Groupe A_4	Groupe B_2	34	91.37%										
35	Groupe A_4	Groupe B_2	35	43.32%										
36	Groupe A_3	Groupe B_2	36	37.80%										
37	Groupe A_4	Groupe B_1	37	26.78%										
38	Groupe A_2	Groupe B_2	38	15.69%										
39	Groupe A_2	Groupe B_4	39	3.24%										
40	Groupe A_4	Groupe B_3	40	95.95%										
41	Groupe A_4	Groupe B_2	41	33.70%										
42	Groupe A_2	Groupe B_2	42	90.86%										
43	Groupe A_2	Groupe B_2	43	90.59%										



Echantillons « appariés » vs. « indépendants »

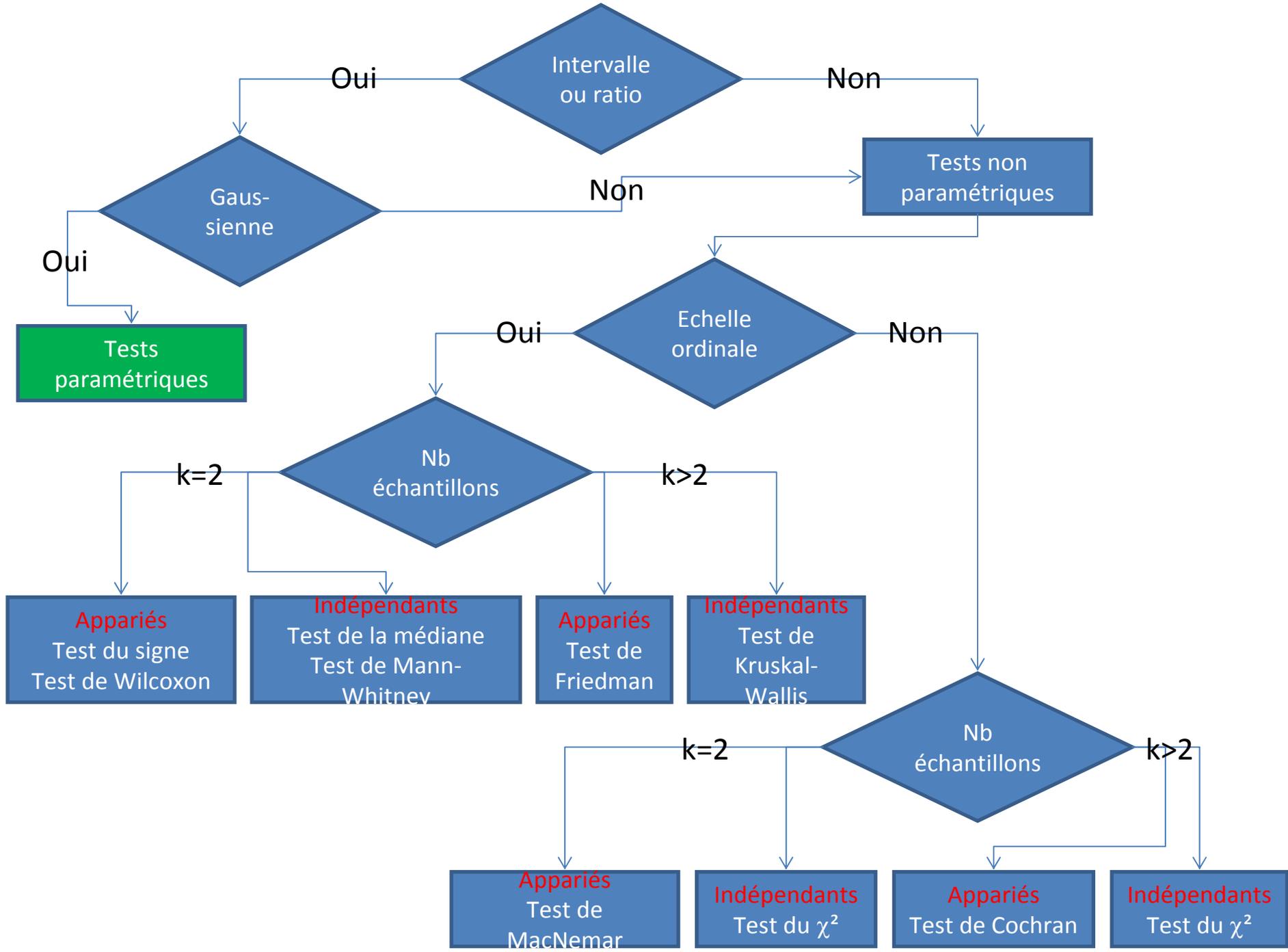
Echantillons « Non indépendants » ou « appariés »

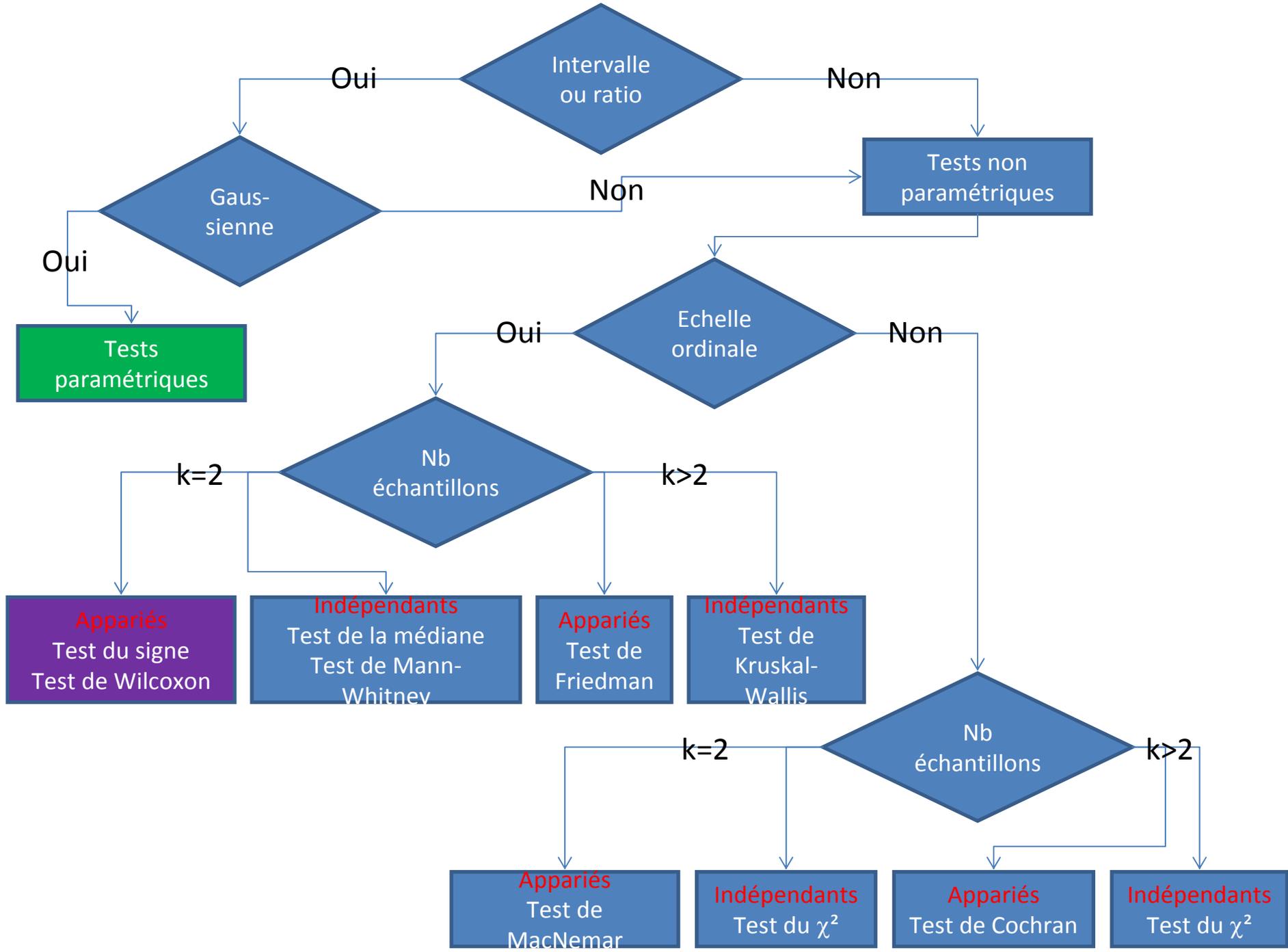
Supposons par exemple qu'une étude pharmacologique porte sur l'efficacité d'un nouveau médicament contre l'hypertension. On mesure dans un premier temps la tension artérielle des patients (« individus ») d'un groupe d'essai. Les valeurs obtenues constituent le premier échantillon S1. Après administration du traitement, la tension artérielle des **mêmes patients** est à nouveau contrôlée, et les valeurs correspondantes sont regroupées dans un second échantillon S2.

Echantillons indépendants

Plusieurs échantillons sont dits indépendants s'ils proviennent de tirages indépendants dans une ou plusieurs populations.

http://www.aiaccess.net/French/Glossaires/GlosMod/Plus/f_gm_plus_echantillons.htm





Echantillons Appariés - Test des signes

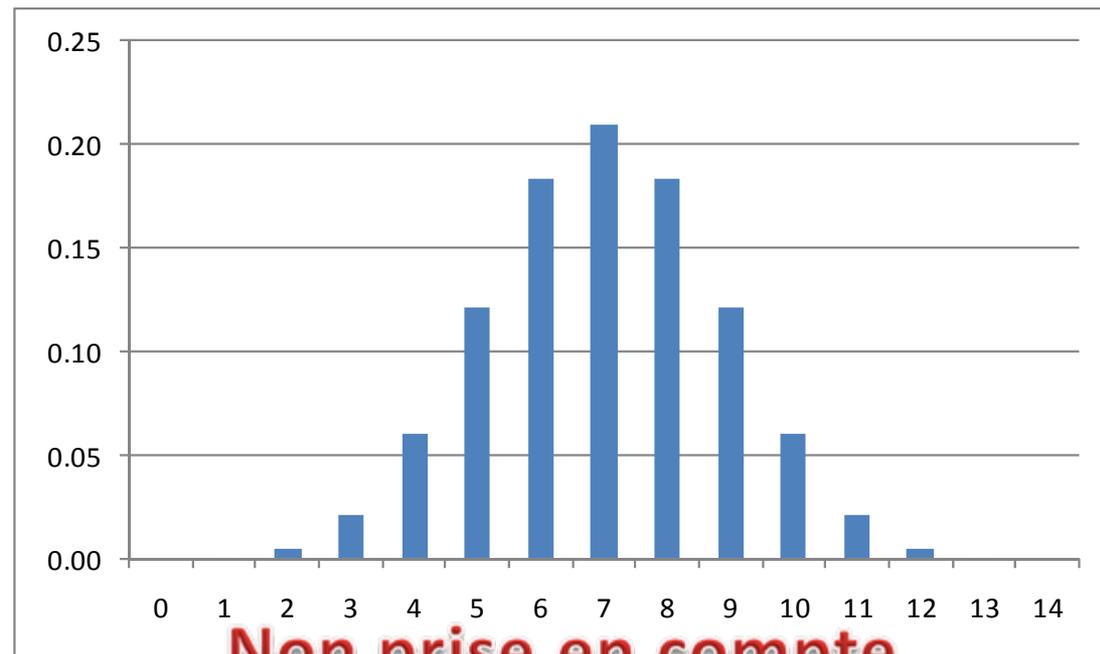
Objectifs

Etablir si deux traitements sont différents ou si un traitement est meilleur que l'autre (variable au moins ordinale)

Hypothèse nulle

H_0 : il n'y a pas de différences entre les deux traitements

Individus	Cas 1	Cas 2
1	20	40
2	18	25
3	24	38
4	14	27
5	5	31
6	26	21
7	15	32
8	29	38
9	15	25
10	9	18
11	25	32
12	31	28
13	35	33
14	12	29



**Non prise en compte
de l'amplitude
des différences**

Echantillons Appariés - Test de Wilcoxon

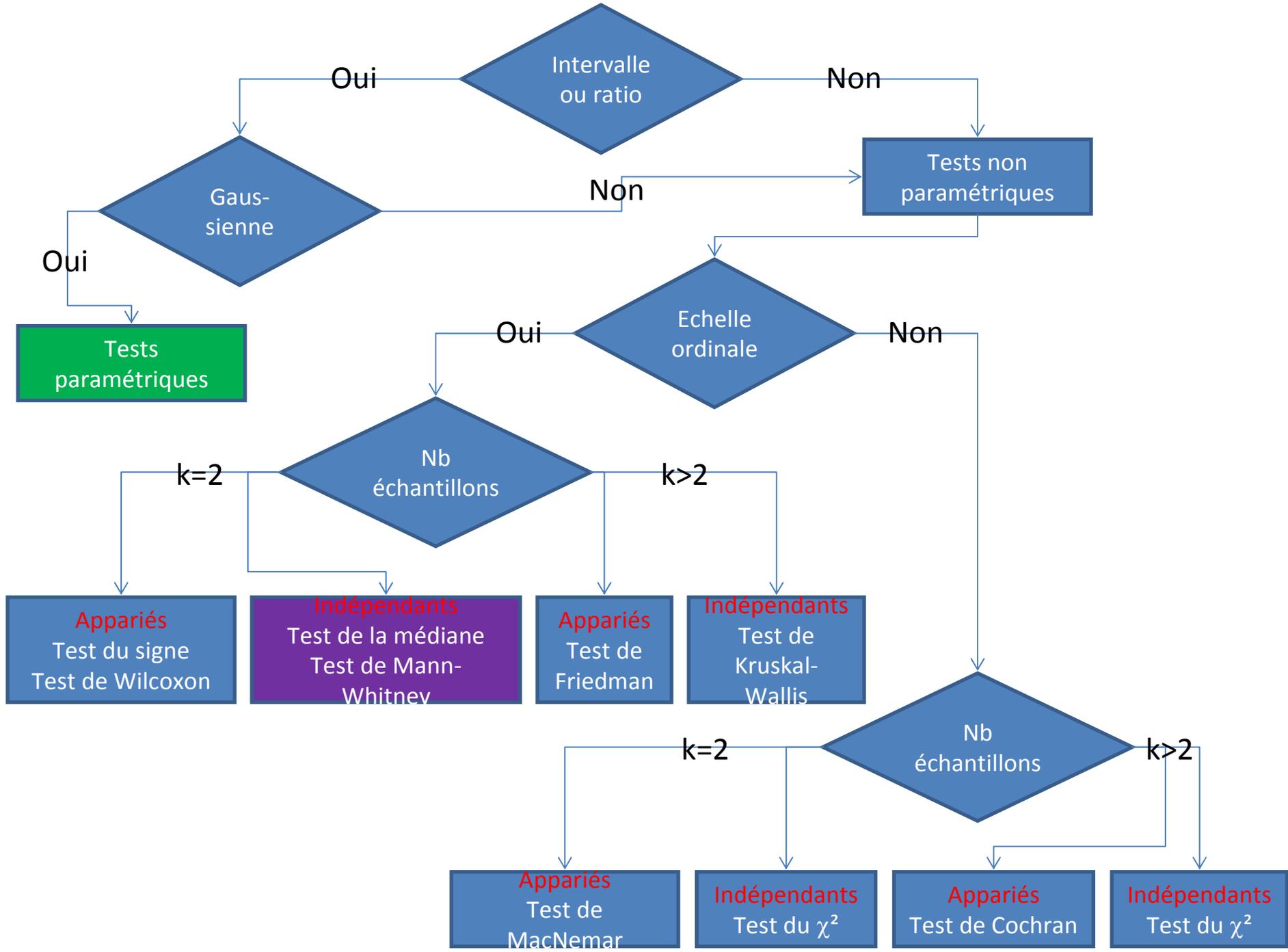
(Wilcoxon Signed-Ranks Test)

Objectifs

Etablir si deux traitements sont différents ou si un traitement est meilleur que l'autre (variable au moins ordinale) tout en tenant compte (dans une certaine mesure) de l'amplitude des différences

Hypothèse nulle

H_0 : il n'y a pas de différences entre les deux traitements



Echantillons Indépendants - Test de Mann-Whitney

Objectifs

Etablir le degré de séparation (ou de recouvrement) entre deux groupes (Expérimental vs. Contrôle) sur une variable au moins ordinale.

Hypothèse nulle

H_0 : les deux séries de valeurs numériques (ou ordinales) sont issues d'une même distribution (\rightarrow elles ne diffèrent pas systématiquement).

Hypothèse alternative

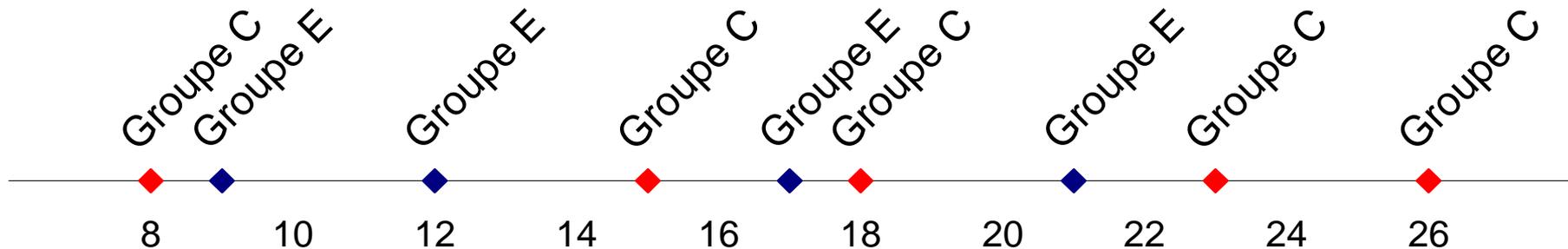
H_A : les deux séries de valeurs numériques (ou ordinales) sont issues de deux distributions différentes (\rightarrow elles diffèrent systématiquement).

NB : avec un test unilatéral, on peut également spécifier le sens de la différence (exemple : le groupe Expérimental a un résultat systématiquement supérieur au groupe de Contrôle).

Echantillons Indépendants - Test de Mann-Whitney

Objectifs

Etablir le degré de séparation (ou de recouvrement) entre deux groupes (Expérimental vs. Contrôle) sur une variable au moins ordinale.



Valeurs	Groupe	Rang	C<E	E<C
8	Groupe C	1	4	
9	Groupe E	2		4
12	Groupe E	3		4
15	Groupe C	4	2	
17	Groupe E	5		3
18	Groupe C	6	1	
21	Groupe E	7		2
23	Groupe C	8	0	
26	Groupe C	9	0	

Echantillons Indépendants - Test de Mann-Whitney

Objectifs

Etablir le degré de séparation (ou de recouvrement) entre deux groupes (Expérimental vs. Contrôle) sur une variable au moins ordinale.

n_1 = nombre de valeurs du groupe 1 (ici $n_1=4$)

n_2 = nombre de valeurs du groupe 2 (ici $n_2=5$)

R_1 = somme des rangs du groupe 1 (ici $R_1=17$)

R_2 = somme des rangs du groupe 2 (ici $R_2=28$)

$$U_1 = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1 = 13$$

$$U_2 = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2 = 7$$

Echantillons Indépendants - Test de la Médiane

Objectifs

Vérifier si deux groupes sont « superposables » au sens où leurs médianes sont identiques (c'est-à-dire qu'elles ont la même valeur que la médiane globale estimée à partir des deux échantillons réunis).

Hypothèse nulle

H_0 : les deux groupes sont séparés dans les mêmes proportions par cette médiane globale.

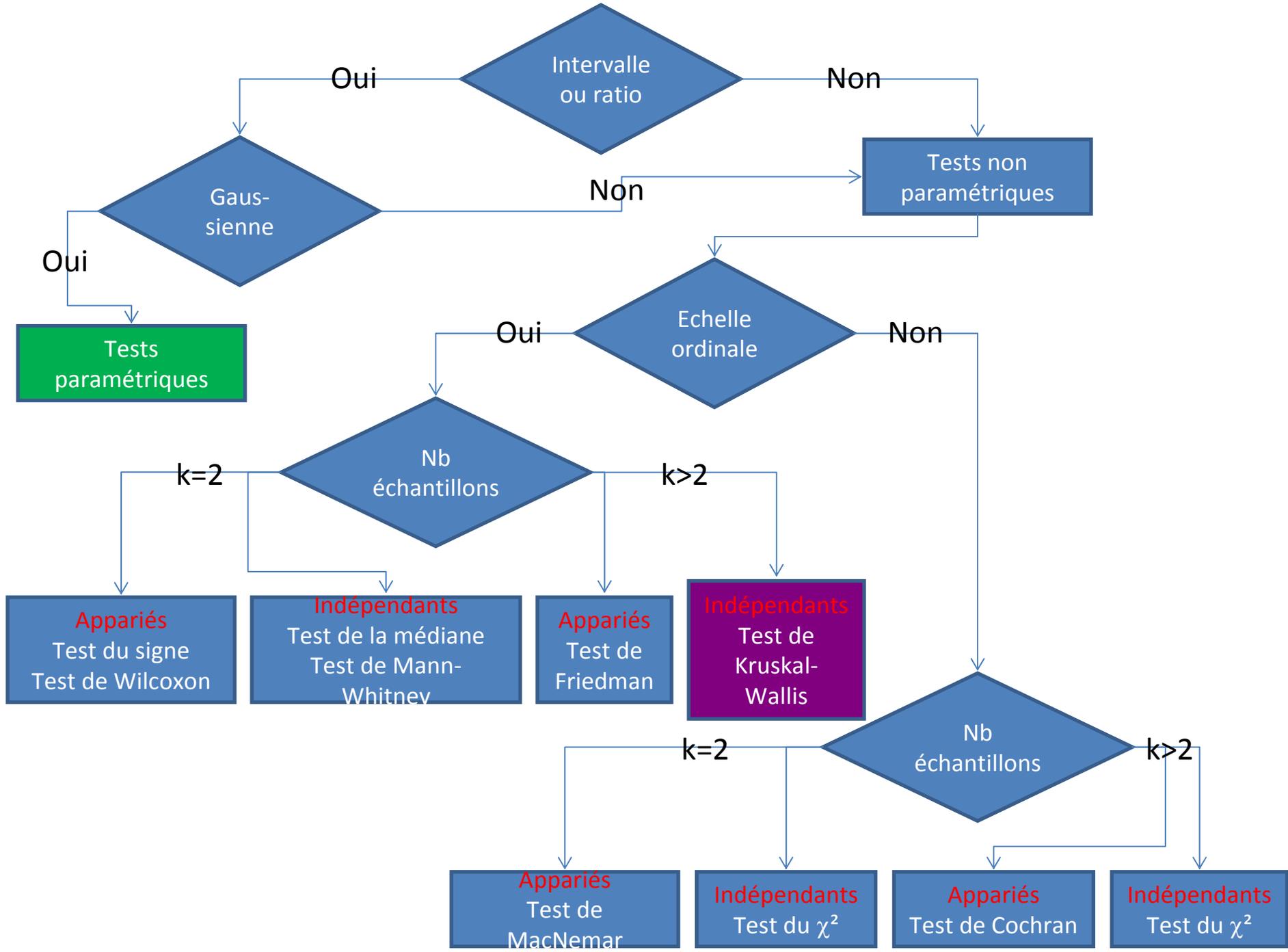
**Test du χ^2 ou
(mieux surtout avec
des petits effectifs)
Test exact de Fisher**

	Groupe E	Groupe C	Total
<= médiane	3	2	5
> médiane	1	3	4
Total	4	5	9

	Groupe E	Groupe C	Total
<= médiane	2.22	2.78	5
> médiane	1.78	2.22	4
Total	4	5	9

	Groupe E	Groupe C	
<= médiane	0.2722	0.2178	
> médiane	0.3403	0.2722	
			1.1025

Fractile 3.84
Proba. Crit. 99.92%



Echantillons Indépendants - Test de Kruskal-Wallis

Objectifs

Etablir le degré de séparation (ou de recouvrement) entre k groupes (indépendants) sur une variable au moins ordinale.

a	b	c	d
68	78	82	54
63	69	73	51
58	58	67	32
51	57	66	74
41	53	61	65

$$H = \left[\frac{12}{n(n+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} \right] - 3(n+1) \approx \chi^2(k-1)$$