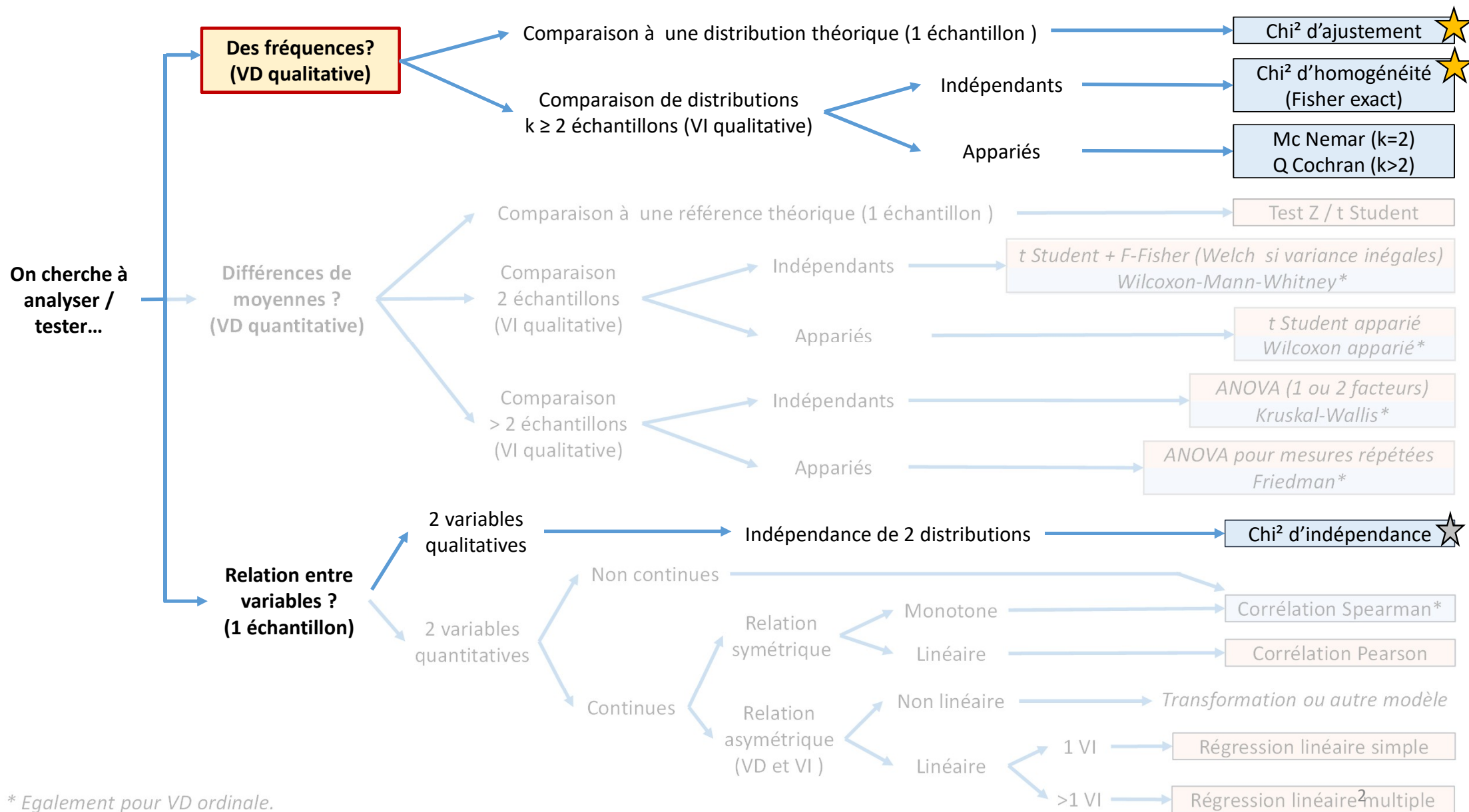


# Biostatistiques

## Partie 3

[diane.zarzoso.lacoste@u-picardie.fr](mailto:diane.zarzoso.lacoste@u-picardie.fr)





# ***Plan du cours***

## **Partie 3. Les tests relevant de la statistique du $\chi^2$**

1. Généralités
2. Test conformité/d'ajustement/d'adéquation
3. Test d'homogénéité /d'égalité
4. Test d'indépendance
5. Cas particuliers



## Tests relevant de la statistique du $\chi^2$

### 1. Généralités

Ensemble de tests **non-paramétrique** et **unilatéraux** visant à comparer des distributions à l'aide d'**EFFECTIFS** (fréquences absolues).

#### Objectif global :

Evaluer si les effectifs relatifs à une ou plusieurs VA qualitatives (ou quantitatives classées), observés sur un ou plusieurs échantillons, sont du même ordre de grandeur que les effectifs moyens théoriques attendus sous l'hypothèse nulle  $H_0$ .

## Tests relevant de la statistique du $\chi^2$

### Principe :

→ Calcul d'une **distance** (écart) entre distributions théorique et observée :  $\chi^2$

### *La question posée :*

Est-ce que cette distance est suffisamment faible pour être due au hasard de l'échantillonnage (→ **H0**)?

Ou est-elle trop importante pour que le hasard à lui seul puisse l'expliquer (→ **H1** : différence significative/réelle) ?

## Tests relevant de la statistique du $\chi^2$

### 3 tests, 3 buts différents:

- **Test conformité/d'ajustement/d'adéquation** : comparer 1 distribution observée à 1 distribution théorique de référence
- **Test d'homogénéité /d'égalité** : comparer des distributions observées entre échantillons indépendants (1 variable, plusieurs échantillons)
- **Test d'indépendance** : tester l'existence d'une liaison entre 2 VA qualitatives issues du même échantillon

## Tests relevant de la statistique du $\chi^2$

### Conditions d'application:

- Les modalités d'une variable sont mutuellement exclusives.
- L'échantillon doit être de grande taille ( $n \geq 50$ ).
- La taille des effectifs théoriques doit être strictement identique à celle des effectifs observés (c.a.d.  $n$ , effectif total).
- Les effectifs théoriques doivent être  $\geq 5$  (ou effectifs théoriques non nuls avec au moins 80%  $\geq 5$ )  $\rightarrow$  sinon regrouper classes adjacentes.

## Tests relevant de la statistique du $\chi^2$

### 2. Test conformité/d'ajustement/d'adéquation

**But :** Tester l'adéquation entre une distribution observée et une distribution de référence\* dont la loi est connue a priori (*anglais : goodness of fit*).

*\* population de référence, loi de probabilités, loi de la génétique ...*

*Ex : Est-ce que la distribution des génotypes observés pour un locus est conforme à la distribution attendue sous le modèle Hardy-Weinberg ?*

**Les données :** effectifs observés pour chaque modalité d'1 variable qualitative (ou quantitative en classes) → comparer avec effectifs théoriques (référence)



## Tests relevant de la statistique du $\chi^2$

### Les hypothèses :

- **H0** : la distribution obs. ne diffère pas significativement de la distribution théo.

→ *L'échantillon est issu de la distribution de référence (population ,loi..)*

- **H1** : la distribution obs. diffère significativement de la distribution théo.

→ *L'échantillon est issu d'une autre loi/population que celle de référence*

## Tests relevant de la statistique du $\chi^2$

**La statistique de test :**

$$\chi^2 = \sum_i \frac{(O_i - C_i)^2}{C_i}$$

Avec  $O_i$  = effectif observé pour modalité « i »,  $C_i$  = effectif théorique attendu

***Comment calculer l'effectif théorique attendu  $C_i$  s'il est inconnu?***

$$C_i = n * p_i$$

Avec  $n$  = effectif total, et  $p_i$  = probabilités théoriques (proportion attendues)

	Facteur	Modalité 1	Modalité 2	Modalité 3	Totaux
<b>Effectifs observés</b>	$O_i$	$O_1$	$O_2$	$O_3$	$n$
<b>Proportions attendues</b>	$p_i$	$p_1$	$p_2$	$p_3$	$1$
<b>Effectifs attendus</b>	$C_i = n * p_i$	$C_1 = n * p_1$	$C_2 = n * p_2$	$C_3 = n * p_3$	$n$