

Statistiques avec



M2 Sciences du Langage

Remi.lafitte@univ-grenoble-alpes.fr

2023-2024

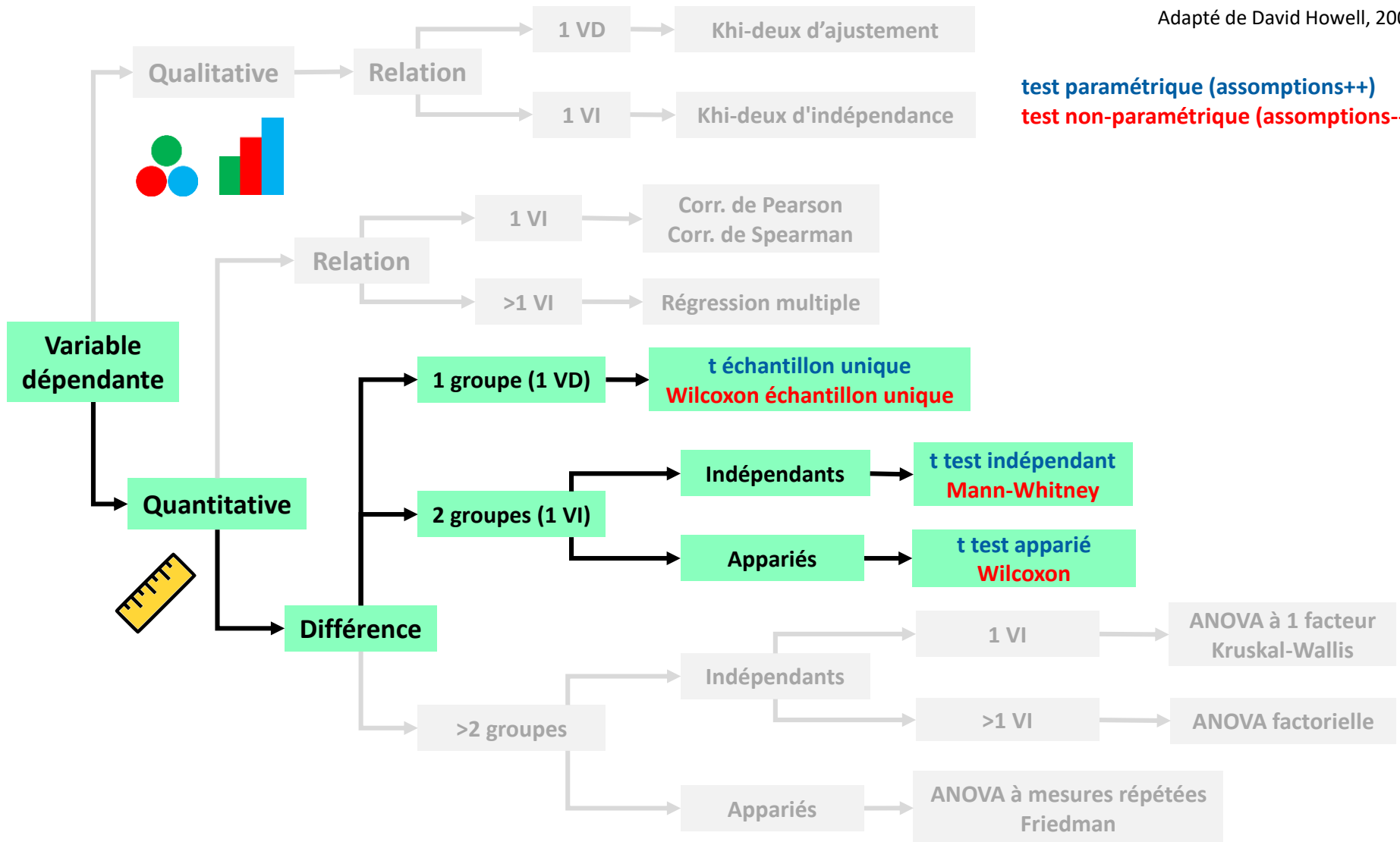
Statistiques descriptives et inférentielles



test t ou t de Student

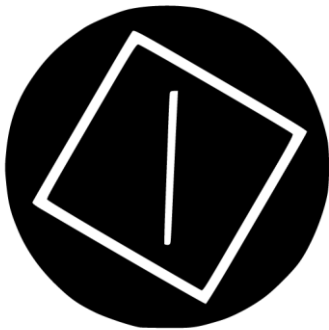
Adapté de David Howell, 2008

test paramétrique (assumptions++)
test non-paramétrique (assumptions--)



test t ou t de Student

- **Dataframe**
- **pp** : participant
- **iq** : QI
- **age**
- **field** : champ d'étude
- **vs_t1 et vs_t2** : mesure de verticale subjective avec un cadre à un temps 1 puis à un temps 2



un sujet lambda a tendance à aligner sa verticale subjective avec l'orientation du carré (penché à +18°)

voir les travaux de Witkin et Asch

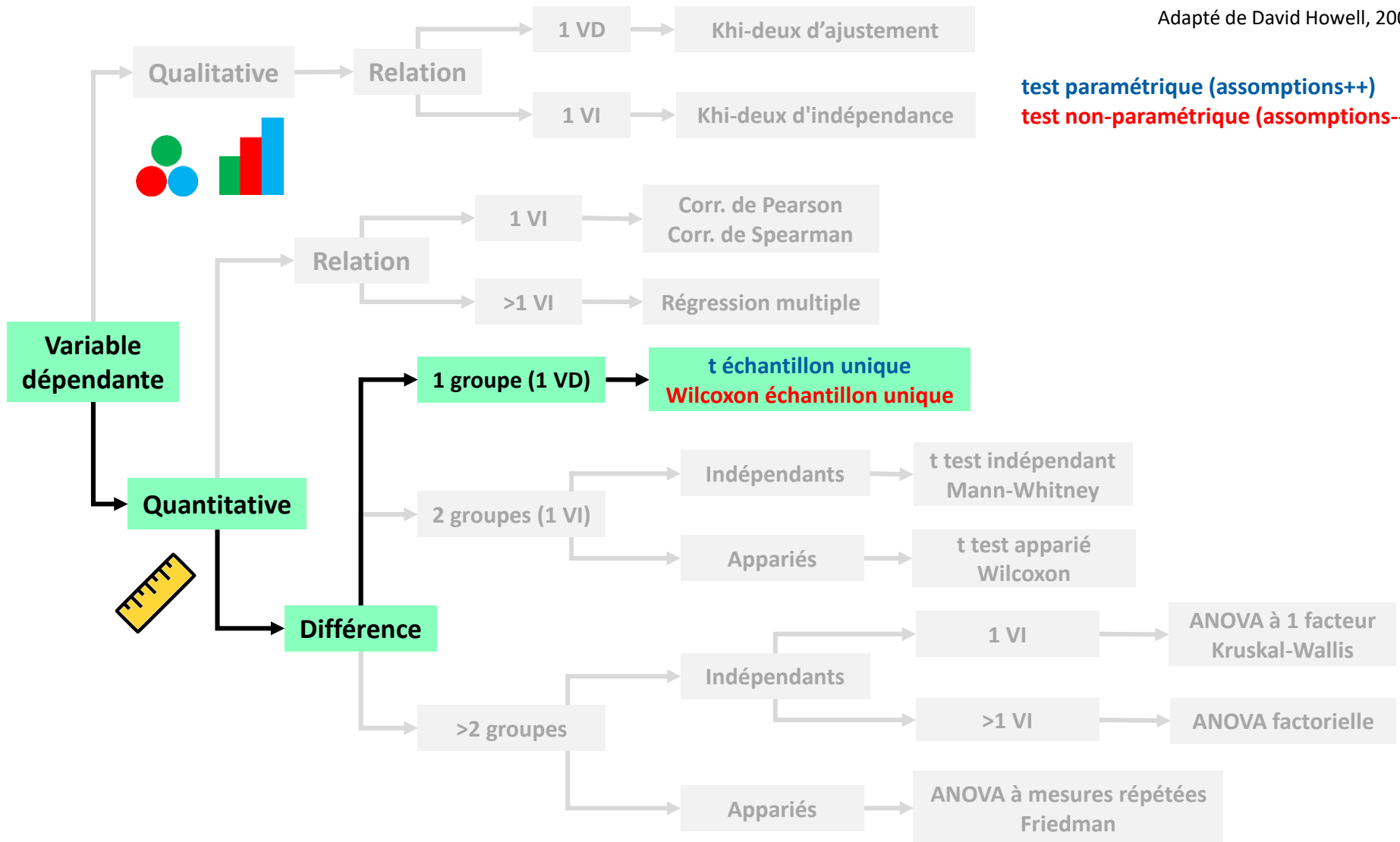
```
DF <- readxl::read_xlsx("perception.xlsx")
DF # montre le DF
```

##	pp	iq	age	field	vs_t1	vs_t2
##	<dbl>	<dbl>	<dbl>	<chr>	<dbl>	<dbl>
## 1	1	110	26	psycho	9.9	13
## 2	2	89	30	psycho	0.3	8.6
## 3	3	120	29	psycho	13.1	0.5
## 4	4	102	31	psycho	12.5	-0.9
## 5	5	113	30	psycho	16.6	1.6
## 6	6	86	23	psycho	6.9	0
## 7	7	108	19	psycho	10.6	6.3
## 8	8	114	32	psycho	0.7	6.7
## 9	9	91	24	psycho	0.6	12.1
## 10	10	81	29	psycho	13.8	14.6
## 11	11	116	30	philos	20.9	-4.2
## 12	12	114	24	philos	22.4	19.6
## 13	13	110	29	philos	20.4	22
## 14	14	120	23	philos	21.5	20.5
## 15	15	110	27	philos	20.4	21.5
## 16	16	118	75	philos	17	20.5
## 17	17	113	23	philos	19.6	22.1
## 18	18	110	22	philos	22.9	18.4
## 19	19	105	21	philos	19.4	17
## 20	20	88	28	philos	22.4	21.5

test t - échantillon unique

Adapté de David Howell, 2008

test paramétrique (assumptions++)
test non-paramétrique (assumptions--)



test t - échantillon unique

- Contexte
- 1 VD quantitative (ordinaire/rapport/intervalle)
- H_0 : moyenne **ECHANTILLON** = moyenne **POPULATION**
- H_1 : moyenne **ECHANTILLON** \neq moyenne **POPULATION**



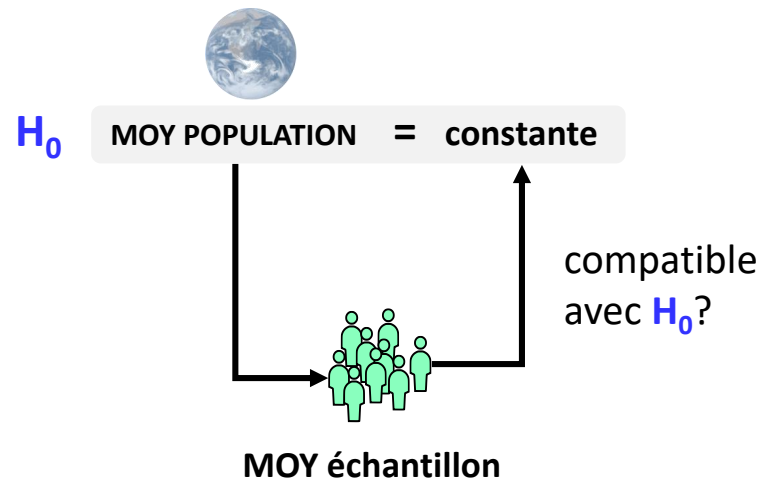
valeur connue OU
supposée




ce test est généralement
utilisé pour savoir si mon
échantillon vient de telle
population

Exemples :

- taille moy > 177 ?
- note < 10 ?
- verticale subjective $\neq 3^\circ$?



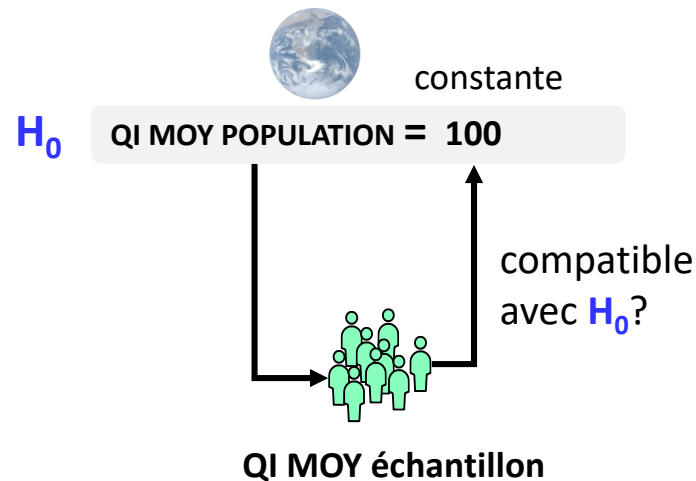
test t - échantillon unique

- Contexte 
- 1 VD quantitative (ordinaire/rapport/intervalle)
- H_0 : moyenne **ECHANTILLON** = moyenne **POPULATION**
- H_1 : moyenne **ECHANTILLON** \neq moyenne **POPULATION**



ce test est généralement utilisé pour savoir si mon échantillon vient de telle population

Exemple ICI : nous avons un groupe de 20 sujets ; nous voulons vérifier si leur **QI** moyen est différent de **100** ou non.



test t - échantillon unique

- Statistiques descriptives

```
# indice de tendances centrales (moyenne et médiane)
summary(DF$iq)

##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   81.00   99.25  110.00  105.90  114.00  120.00

# indice de dispersion (écart type et quantiles)
sd(DF$iq)

## [1] 12.16077

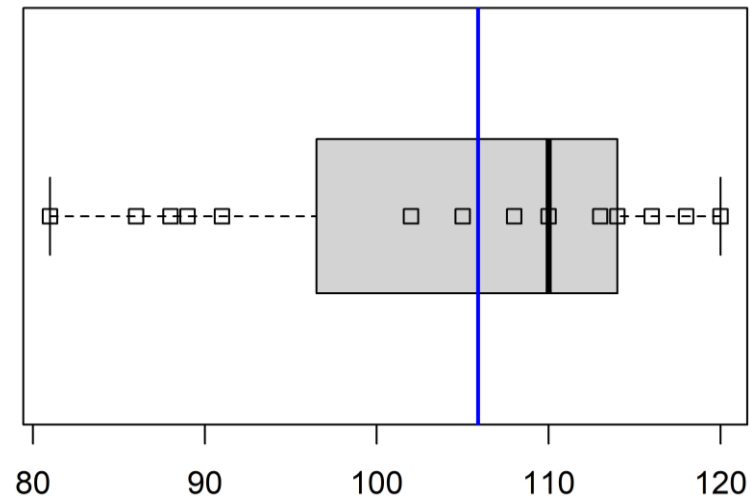
quantile(x = DF$iq, probs = c(0.25, 0.75))

##      25%      75%
##   99.25  114.00
```


test t - échantillon unique

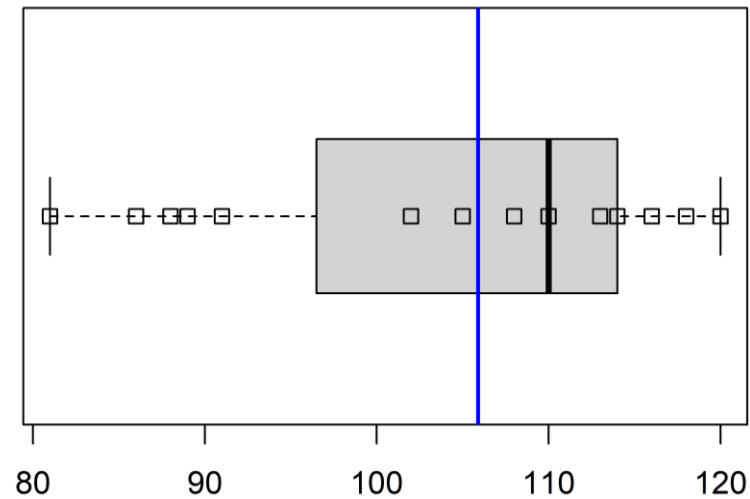
- Statistiques descriptives

```
# graphique  
boxplot(DF$iq, horizontal = T)  
graphics::stripchart(DF$iq, add = T) # rajoute les observations individuelles  
abline(v = mean(DF$iq), col = "blue", lwd = 2) # plot la moyenne
```



test t - échantillon unique

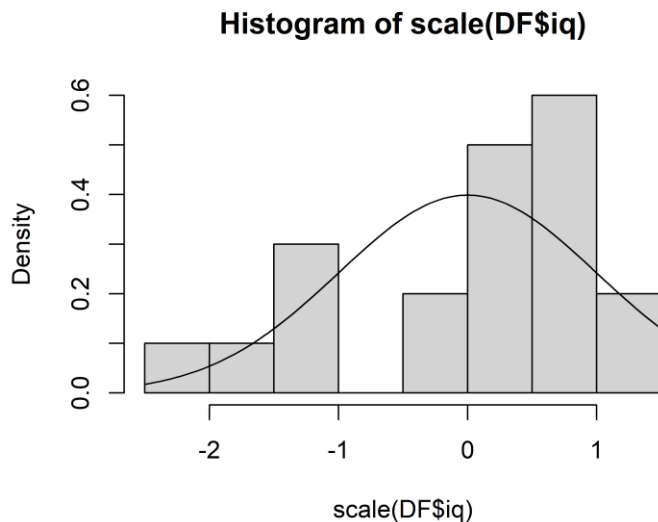
- Conditions d'application
- (1) Outlier ? RAS



test t - échantillon unique

- Conditions d'application
- (2) Normalité de la distribution ? mitigée...

```
# histogramme de densité sur la variable centrée-réduite  
hist(scale(DF$iq), prob=T); curve(dnorm(x), add=T)
```



```
# test de Shapiro Wilk  
shapiro.test(DF$iq)  
  
##  
##  Shapiro-Wilk normality test  
##  
## data:  DF$iq  
## W = 0.86657, p-value = 0.01023
```

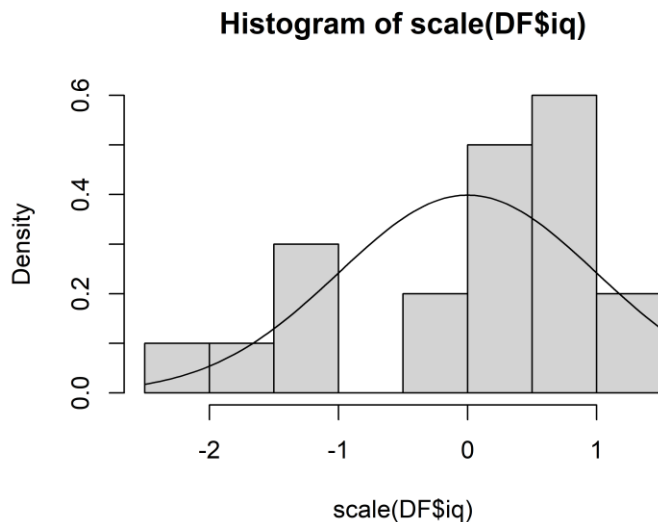
test t - échantillon unique

- Conditions d'application
- (2) Normalité de la distribution ? mitigée...

- Si distribution normale, on applique le test sur la **moyenne**
- Si distribution anormale, on applique le test sur la **médiane**

test t

test de Wilcoxon



test t - échantillon unique

- Statistiques inférentielles

```
t.test(x= DF$iq, mu = 100)

##
## One Sample t-test
##
## data: DF$iq
## t = 2.1697, df = 19, p-value = 0.04292
## alternative hypothesis: true mean is not equal to 100
## 95 percent confidence interval:
##  100.2086 111.5914
## sample estimates:
## mean of x
##      105.9
```

MOY population

n - 1

MOY échantillon

test t - échantillon unique

- Statistiques inférentielles

MED population

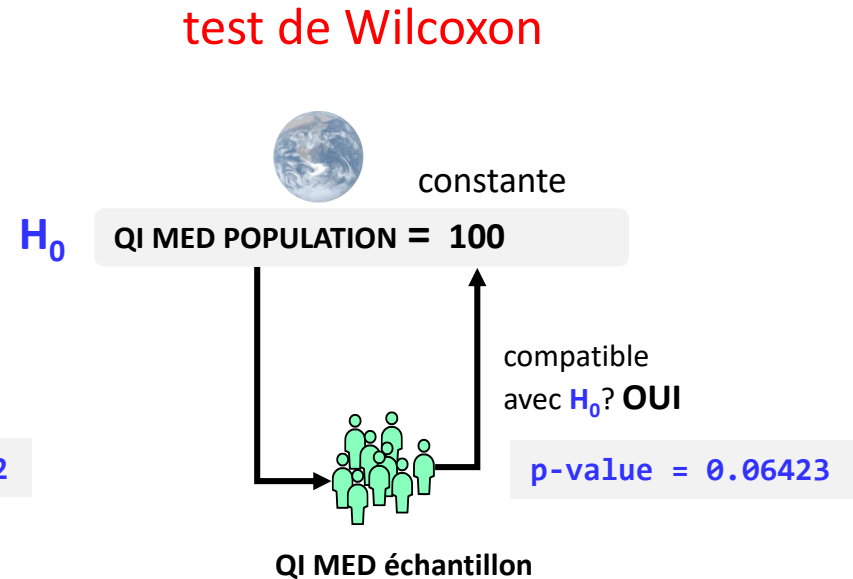
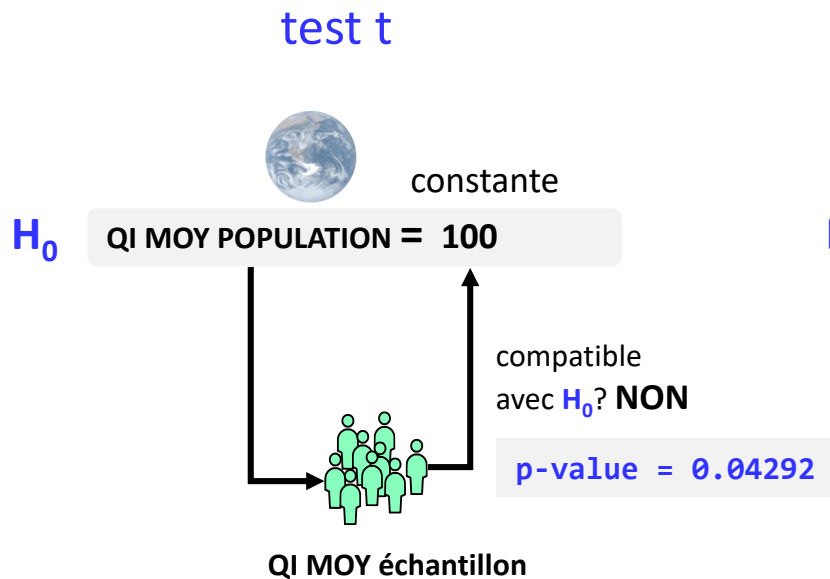


```
wilcox.test(x= DF$iq, mu = 100)

##
##  Wilcoxon signed rank test with continuity correction
##
## data:  DF$iq
## V = 155, p-value = 0.06423
## alternative hypothesis: true location is not equal to 100
```

test t - échantillon unique

- Statistiques inférentielles



- comme la normalité des données est douteuse ici, nous faisons plutôt confiance au test de Wilcoxon et on ne rejette pas H_0 . On ne peut pas rejeter l'hypothèse que notre échantillon provienne d'une population "normale".
- Cet exemple nous montre qu'il est très facile de faire du *p-hacking* !

test t - échantillon unique

- Rédaction

A one sample Wilcoxon test was performed to evaluate whether there was a difference between the IQ of students and the IQ of the general population (i.e., 100). The median IQ in students **(110 [99.25; 114])** was **not significantly** different from **100**, $p = .06$.

mediane [Q1; Q3]

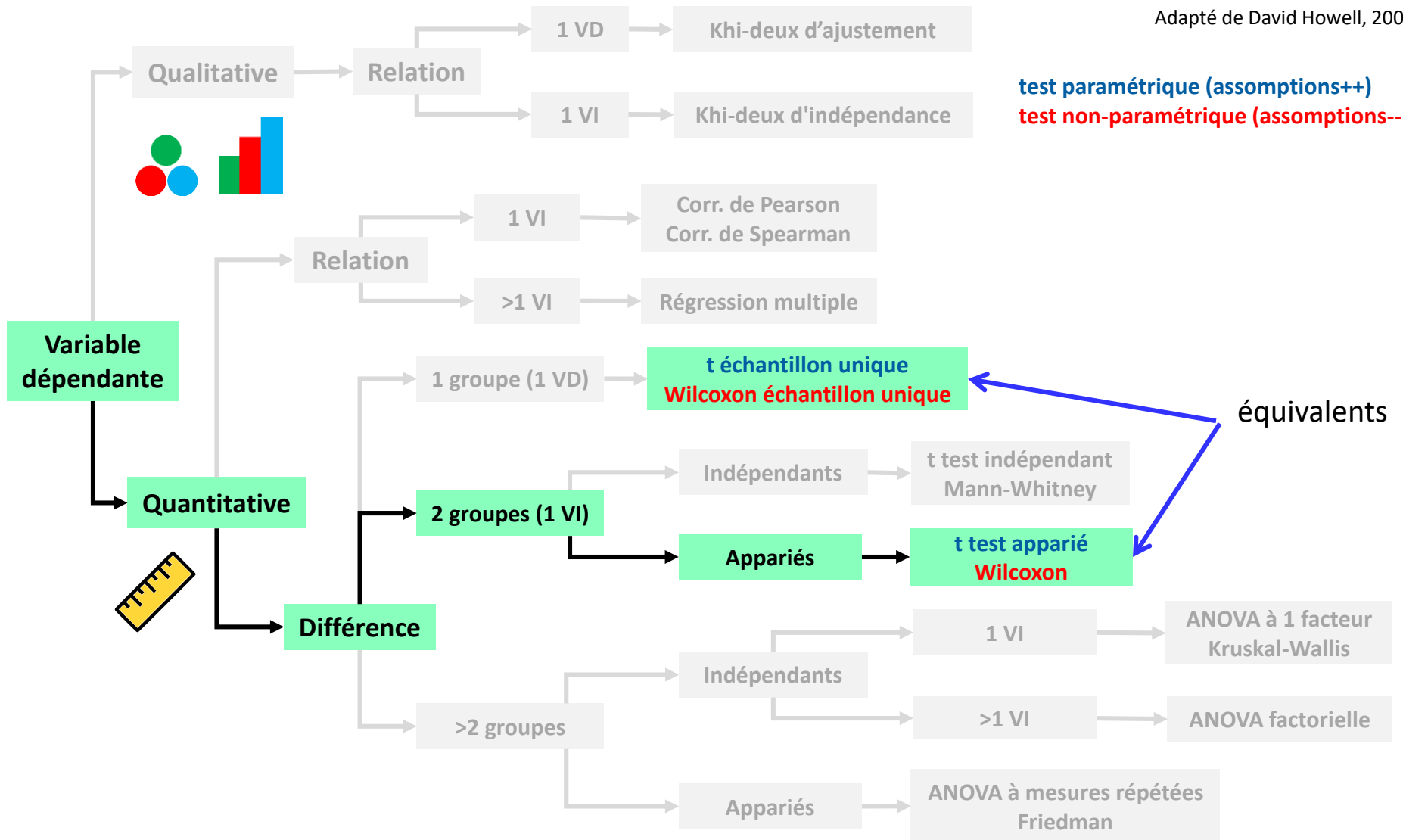


A one sample t-test was performed to evaluate whether there was a difference between the IQ of students and the IQ of the general population (i.e., 100). The mean IQ in students **(105.9 ± 12.2)** differed **significantly** from **100**, $t(19) = 2.17$, $p = .04$.

moy ± écart-type



test paramétrique (assumptions++)
test non-paramétrique (assumptions--)



test t – échantillons appariés

- Contexte




- 1 VD quantitative (ordinaire/rapport/intervalle)
- 1 VI qualitative **intra-sujet** à deux modalités

pp	vs_t1	vs_t2
<dbl>	<dbl>	<dbl>
1	9.9	13
2	0.3	8.6
3	13.1	0.5
4	12.5	-0.9



Exemple de VI intra-sujet : la VI
temps. **TOUS les SUJETS** passent par
les modalités "vs_**t1**" ET "vs_**t2**"

test t – échantillons appariés

- Contexte 
- 1 VD quantitative (ordinaire/rapport/intervalle)
- 1 VI qualitative **intra-sujet** à deux modalités

modalités sur la même LIGNE



Pour les VI INTRA-sujets on privilégie le format COURT

pp	vs_t1	vs_t2
<dbl>	<dbl>	<dbl>
1	9.9	13
2	0.3	8.6
3	13.1	0.5
4	12.5	-0.9

Exemple de VI intra-sujet : la VI
temps. TOUS les SUJETS passent par
 les modalités "vs_t1" ET "vs_t2"

- 



pp	vs_t1	vs_t2
<dbl>	<dbl>	<dbl>
1	9.9	13
2	0.3	8.6
3	13.1	0.5
4	12.5	-0.9

*modalités sur la
même COLONNE*

[illegible]

test t – échantillons appariés

- Contexte



- 1 VD quantitative (ordinaire/rapport/intervalle)
- 1 VI qualitative **intra-sujet** à deux modalités

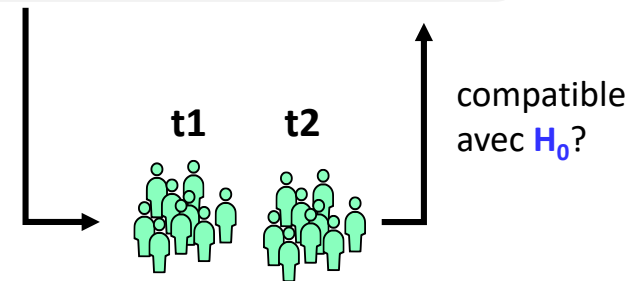
Q? : est-ce que la verticale subjective ("vs") **change** entre t1 et t2 ?

- H_0 : différence de MOYENNES entre temps 1 et temps 2 = 0
- H_1 : différence de MOYENNES entre temps 1 et temps 2 $\neq 0$

POPULATION



$$\text{MOY "vs_t2" (-) MOY "vs_t1"} = 0^\circ$$



test t – échantillons appariés

- Statistiques descriptives
 - On est intéressé par le **changement de vs entre t1 et t2**

```
DF$vs_change = DF$vs_t2 - DF$vs_t1  
# nb : (t2 - t1) est plus facile à interpreter que l'inverse
```

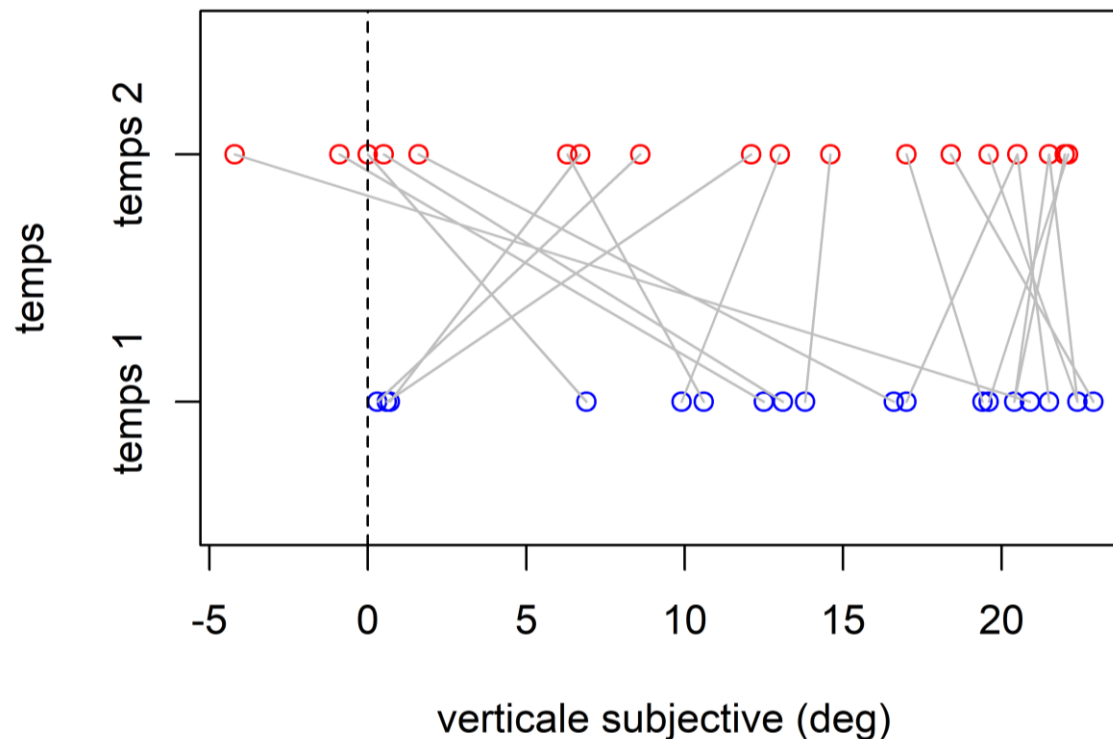
- Le code est le même que précédemment

```
# indice de tendances centrales (moyenne et médiane)  
summary(DF$vs_change)  
# indice de dispersion (écart type et quantiles)  
sd(DF$vs_change)  
quantile(x = DF$vs_change, probs = c(0.25, 0.75))  
# graphique  
boxplot(DF$vs_change, horizontal = T)  
graphics::stripchart(DF$vs_change, add = T) # rajoute les observations individuelles  
abline(v = mean(DF$vs_change), col = "blue", lwd = 2)  
# plot la moyenne
```

test t – échantillons appariés

- Statistiques descriptives

Intéressant de représenter graphiquement l'**évolution entre t1 et t2**
pour chaque participant



ici un trait = un sujet

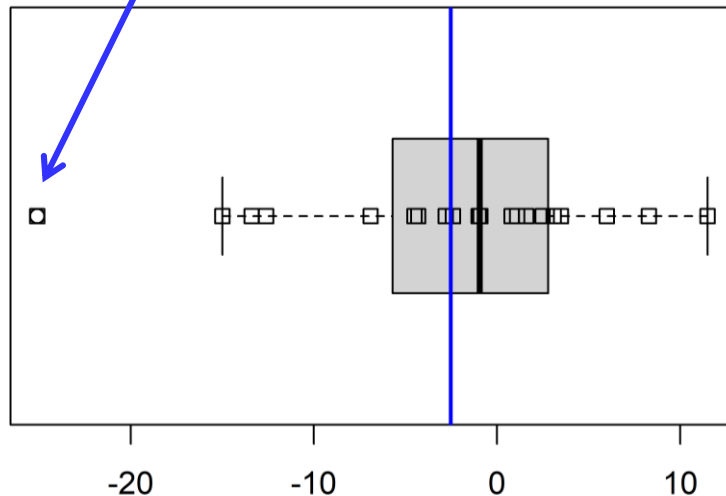
*voir le code,
relativement complexe,
sur Rmd*

test t – échantillons appariés

- Conditions d'application

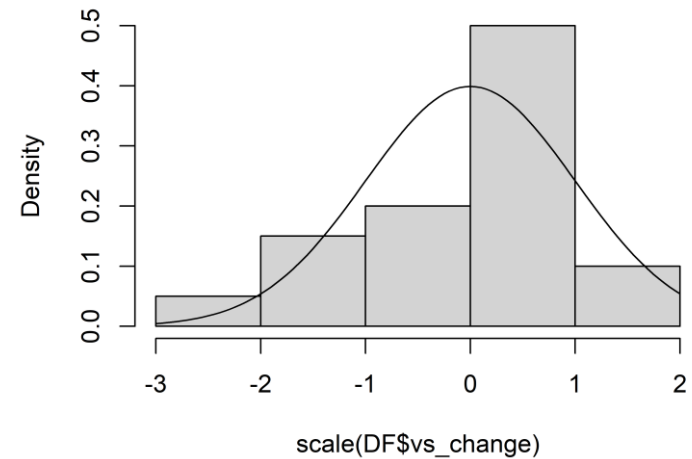
Outlier : 1 suspect

Normalité : bonne



```
## Shapiro-Wilk normality test
##
## data: DF$vs_change
## W = 0.94074, p-value = 0.2476
```

Histogram of scale(DF\$vs_change)



test t – échantillons appariés

- Statistiques inférentielles

```
t.test(x= DF$vs_change, mu = 0)
##
## One Sample t-test
##
## data: DF$vs_change
## t = -1.2983, df = 19, p-value = 0.2097
## alternative hypothesis: true mean is not
## equal to 0
## 95 percent confidence interval:
## -6.59549 1.54549
## sample estimates:
## mean of x
## -2.525
```

← changement de vs moyen

```
wilcox.test(x= DF$vs_change, mu = 0)
## Wilcoxon signed rank exact test
##
## data: DF$vs_change
## V = 80, p-value = 0.3683
## alternative hypothesis: true locat
## ion is not equal to 0
```

test t – échantillons appariés

- Statistiques inférentielles

```
t.test(x= DF$vs_change, mu = 0)

##
## One Sample t-test
##
## data: DF$vs_change
## t = -1.2983, df = 19, p-value = 0.2097
## alternative hypothesis: true mean is not
## equal to 0
## 95 percent confidence interval:
## -6.59549 1.54549
## sample estimates:
## mean of x
## -2.525
```

← changement de vs moyen

- On ne rejette pas H_0 . On ne généralise pas ce changement de verticale subjective à la population

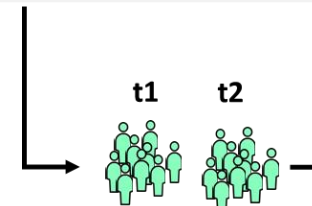
```
wilcox.test(x= DF$vs_change, mu = 0)
## Wilcoxon signed rank exact test
##
## data: DF$vs_change
## V = 80, p-value = 0.3683
## alternative hypothesis: true locat
## ion is not equal to 0
```

POPULATION

H_0



MOY "vs_t2" (-) MOY "vs_t1"
= 0°



compatible
avec H_0 ?

OUI, car $p > .05$

test t – échantillons appariés

- Statistiques inférentielles

BONUS : il est tout de même possible d'effectuer un t-test pour échantillons appariés sur des **données au format long**

- transformation du format court à long avec la fonction **reshape**

```
DF = data.frame(DF) # reshape requiert ABSOLUMENT
ce format
DF_long =
stats::reshape(data = DF,
# notre DF
                direction = "long",
# indique que l'on veut que la VI "temps" (vs_t1,
vs_t2) soit en format long, au lieu d'un format
court.
                varying = c("vs_t1", "vs_t2"),
# Les modalités de la VI
                v.names = "angle")
# Le nom de la VD
```

##	pp	iq	age	field	vs_change	time	angle	id
## 1.1	1	110	26	psycho	3.1	1	9.9	1
## 2.1	2	89	30	psycho	8.3	1	0.3	2
## 3.1	3	120	29	psycho	-12.6	1	13.1	3
## 4.1	4	102	31	psycho	-13.4	1	12.5	4
## 5.1	5	113	30	psycho	-15.0	1	16.6	5
...								
## 16.2	16	118	75	philos	3.5	2	20.5	16
## 17.2	17	113	23	philos	2.5	2	22.1	17
## 18.2	18	110	22	philos	-4.5	2	18.4	18
## 19.2	19	105	21	philos	-2.4	2	17.0	19
## 20.2	20	88	28	philos	-0.9	2	21.5	20

VI VD



un bon tuto sur les formats court et long :

https://rstudio-pubs-static.s3.amazonaws.com/594650_88973f23f57c4d60b346abe0bc38801b.html

test t – échantillons appariés

- Statistiques inférentielles

BONUS : il est tout de même possible d'effectuer un t-test pour échantillons appariés sur des **données au format long**

- la syntaxe des fonctions **t.test** et **wilcox.test** change légèrement :

```

                VD      VI
t.test(formula = angle ~ time,
       data = DF_long,
       paired = TRUE)

wilcox.test(formula = angle ~ time,
            data = DF_long,
            paired = TRUE)

```

/!\ indique que les données sont appariées

						VI	VD	
##	pp	iq	age	field	vs_change	time	angle	id
## 1.1	1	110	26	psycho	3.1	1	9.9	1
## 2.1	2	89	30	psycho	8.3	1	0.3	2
## 3.1	3	120	29	psycho	-12.6	1	13.1	3
## 4.1	4	102	31	psycho	-13.4	1	12.5	4
## 5.1	5	113	30	psycho	-15.0	1	16.6	5
...								
## 16.2	16	118	75	philos	3.5	2	20.5	16
## 17.2	17	113	23	philos	2.5	2	22.1	17
## 18.2	18	110	22	philos	-4.5	2	18.4	18
## 19.2	19	105	21	philos	-2.4	2	17.0	19
## 20.2	20	88	28	philos	-0.9	2	21.5	20

test t – échantillons appariés

- Rédaction

A paired samples *t*-test was performed to evaluate whether the subjective vertical varied between the first and second assessments. The results indicated that the **change** in subjective vertical ($2.53 \pm 8.7^\circ$) **was not significantly different from 0° , $t(19) = 1.30$, $p = .21$.**

A Wilcoxon-signed-rank test was performed to evaluate whether the subjective vertical varied between the first and second assessments. The results indicated that the **change** in subjective vertical ($-0.95 [-5.1; 2.7]^\circ$) **was not significantly different from 0° ($p = .37$, Wilcoxon signed-rank test)**

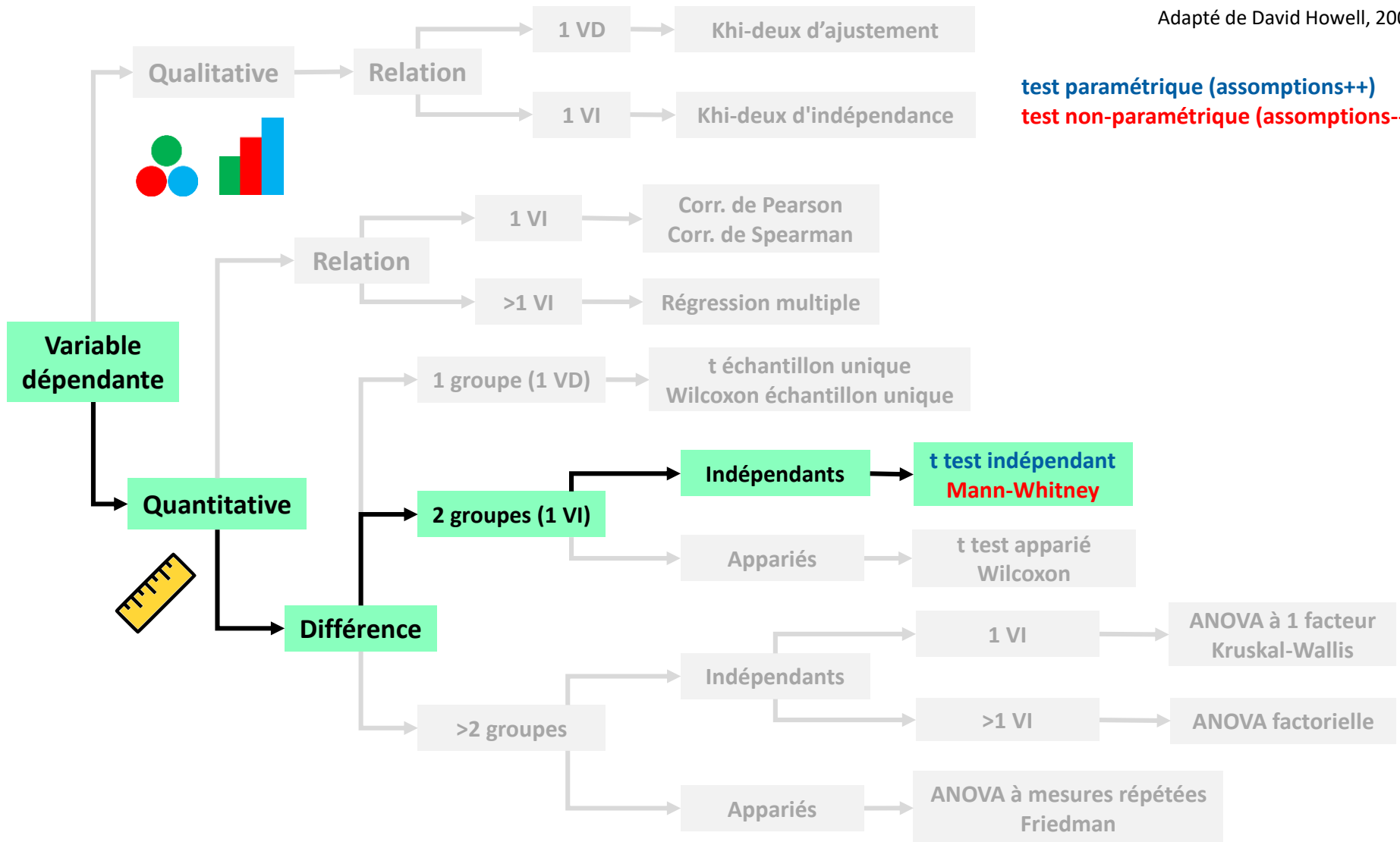


possibilité de mentionner le
nom du test

test t – échantillons appariés

Adapté de David Howell, 2008

test paramétrique (assumptions++)
test non-paramétrique (assumptions--)



test t – échantillons indépendants

- Contexte



- 1 VD quantitative (ordinaire/rapport/intervalle)
- 1 VI qualitative **inter-sujet** à deux modalités

##	pp	iq	age	field	vs_change	time	angle	id
## 1.1	1	110	26	psycho	3.1	1	9.9	1
## 2.1	2	89	30	psycho	8.3	1	0.3	2
## 3.1	3	120	29	psycho	-12.6	1	13.1	3
## 4.1	4	102	31	psycho	-13.4	1	12.5	4
## 5.1	5	113	30	psycho	-15.0	1	16.6	5
...								
## 16.2	16	118	75	philo	3.5	2	20.5	16
## 17.2	17	113	23	philo	2.5	2	22.1	17
## 18.2	18	110	22	philo	-4.5	2	18.4	18
## 19.2	19	105	21	philo	-2.4	2	17.0	19
## 20.2	20	88	28	philo	-0.9	2	21.5	20



Exemple de VI inter-sujet : la VI **field**. Les **SUJETS** ne **PEUVENT PAS** être à la fois en "psycho" et en "philo"

test t – échantillons indépendants

- Contexte

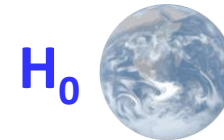


- 1 VD quantitative (ordinaire/rapport/intervalle)
- 1 VI qualitative **inter-sujet** à deux modalités

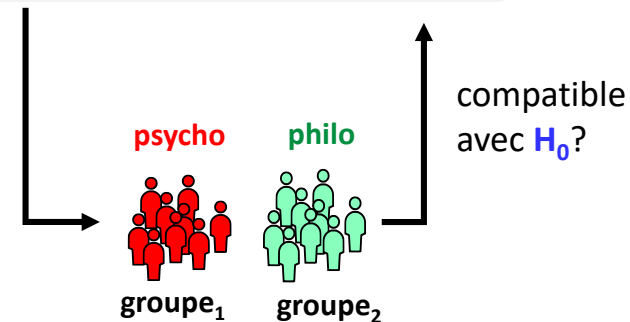
Q? : différence de vs moyennes entre les groupes psycho et philo ?

- H_0 : différence de MOYENNES entre psycho et philo = 0
- H_1 : différence de MOYENNES entre psycho et philo $\neq 0$

POPULATION



MOY "psycho" (-) MOY "philo"
= 0°




test t – échantillons indépendants

- Statistiques descriptives
- Nous ne sommes pas intéressés ici par l'effet du temps sur la vs. Nous allons donc **moyenner "vs_t1" et "vs_t2"**

```
DF$vs_moy = (DF$vs_t1 + DF$vs_t2)/2
```

- Nous calculons les indices de tendance centrale et de dispersion pour chaque groupe de la VI "field" avec la fonction **aggregate**

VD VI FONCTION

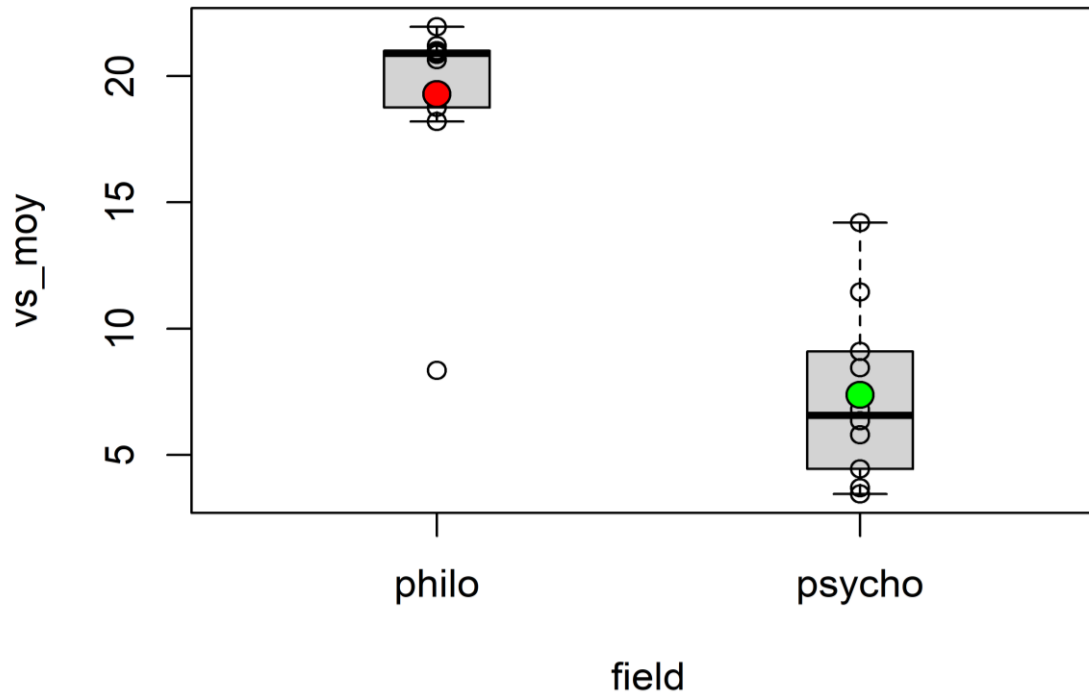


```
stats::aggregate(x = vs_moy ~ field, data = DF, FUN = summary)
stats::aggregate(x = vs_moy ~ field, data = DF, FUN = sd)

##      field    vs_moy
## 1  philo 4.010736
## 2 psycho 3.474450
```

test t – échantillons indépendants

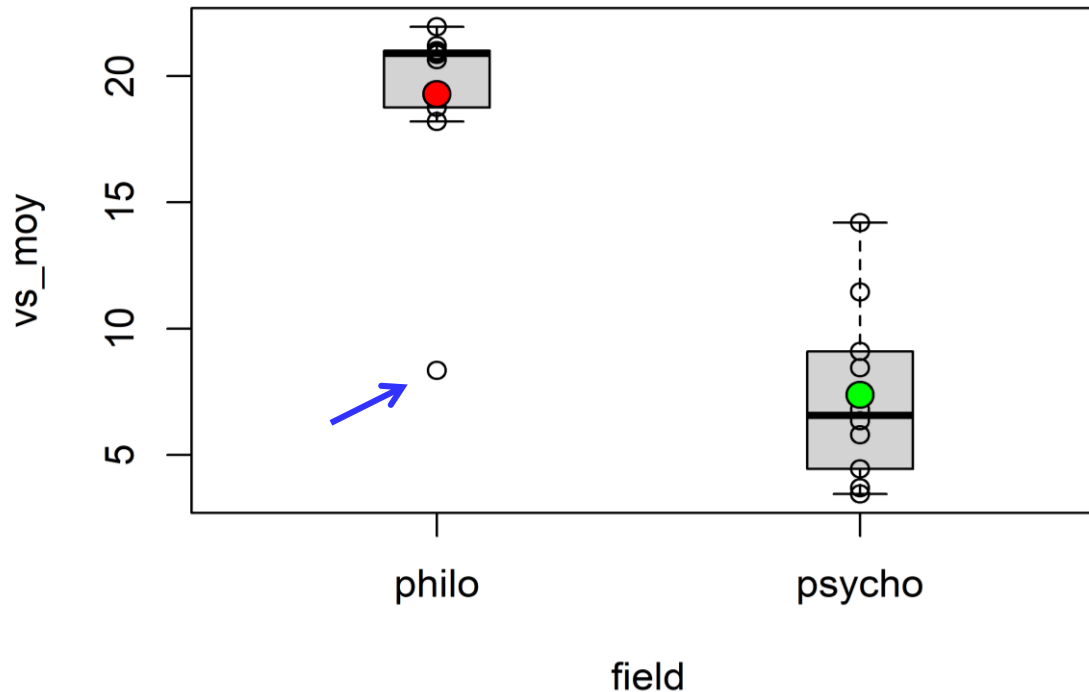
- Statistiques descriptives
- Sur un plan graphique tout est possible ; par ex combiner boxplot et moyenne



voir le Rmd pour le code

test t – échantillons indépendants

- Conditions d'applications
- (1) Outlier à vérifier dans **chaque groupe** : une observation suspecte dans le groupe "philo"



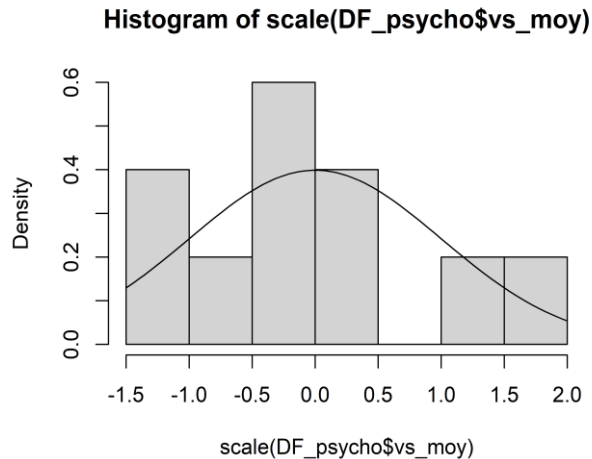
voir le Rmd pour le code

test t – échantillons indépendants

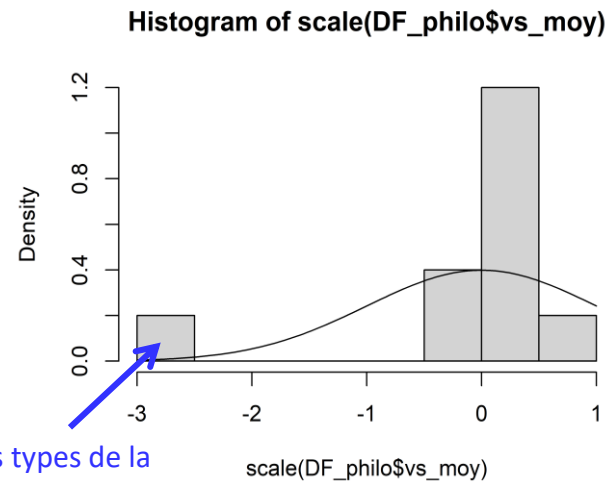
- Conditions d'applications
- (2) Normalité des données pour **chaque groupe**
- mauvaise dans le groupe philo à cause de l'outlier

```
# on sépare notre DF par groupe
DF_psycho = DF[DF$field == "psycho",]
DF_philo = DF[DF$field == "philo",]
```

```
hist(scale(DF_psycho$vs_moy), prob = T)
curve(dnorm(x), add = T)
```



```
hist(scale(DF_philo$vs_moy), prob = T)
curve(dnorm(x), add = T)
```



vs < 3 écarts types de la moyenne du groupe

test t – échantillons indépendants

- Conditions d'applications
- (3) Variances (ou écarts-types) homogènes entre les groupes

```
stats::aggregate(x = vs_moy ~ field, data = DF, FUN = sd)
##      field    vs_moy
## 1  philo  4.010736
## 2 psycho  3.474450
```

- En pratique on ne soucie pas de cette condition d'application, car R utilise par défaut un **test t pour variances inégales** (aussi appelé **Welch t-test**)

test t – échantillons indépendants

- Statistiques inférentielles

VD prédit par VI

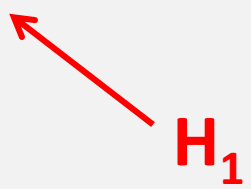
```
t.test(vs_moy ~ field, DF) # teste une différence de moyenne  
##  
## Welch Two Sample t-test  
##  
## data: vs_moy by field  
## t = 7.1006, df = 17.641, p-value = 1.439e-06  
## alternative hypothesis: true difference in means between group philo and group  
## psycho is not equal to 0  
## 95 percent confidence interval:  
## 8.38445 15.44555  
## sample estimates:  
## mean in group philo mean in group psycho  
## 19.290 7.375
```

moyennes

test t – échantillons indépendants

- Statistiques inférentielles

```
t.test(vs_moy ~ field, DF) # teste une différence de moyenne
##
##  Welch Two Sample t-test
##
## data:  vs_moy by field
## t = 7.1006, df = 17.641, p-value = 1.439e-06
## alternative hypothesis: true difference in means between group philo and group
## psycho is not equal to 0
## 95 percent confidence interval:
##    8.38445 15.44555
## sample estimates:
##  mean in group philo mean in group psycho
##           19.290           7.375
```



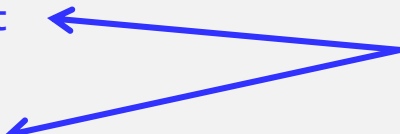
H₁

test t – échantillons indépendants

- Statistiques inférentielles

```
t.test(vs_moy ~ field, DF) # teste une différence de moyenne
##
##  Welch Two Sample t-test
##
## data:  vs_moy by field
## t = 7.1006, df = 17.641, p-value = 1.439e-06
## alternative hypothesis: true difference in means between group philo and group
## psycho is not equal to 0
## 95 percent confidence interval:
##    8.38445 15.44555
## sample estimates:
## mean in group philo mean in group psycho
##           19.290           7.375
```

test t pour variances inégales



On avait 10^{-6} % de chances d'avoir une différence de moyennes au moins aussi extrême sous H_0 . On rejette H_0 .

test t – échantillons indépendants

- Statistiques inférentielles
- Comme nous avons un outlier, on lance aussi le test non-paramétrique de **Mann-Whitney**, dit aussi **Wilcoxon-Mann-Whitney** (d'où le nom de la fonction R : "wilcox.test")

```
wilcox.test(vs_moy ~ field, DF) # teste une différence de médiane
##
## Wilcoxon rank sum test with continuity correction
##
## data: vs_moy by field
## W = 96, p-value = 0.0005801
## alternative hypothesis: true location shift is not equal to 0
```

Bien que la p-valeur soit plus conservatrice, la conclusion est identique.

test t – échantillons indépendants

- Rédaction

possibilité de présenter les résultats comme ceci (M = mean, SD = standard deviation)

The subjective vertical of the 10 students in philosophy (***M* = 19.3°, *SD* = 4.01**), compared to that of the 10 students in psychology (***M* = 6.6°, *SD* = 3.47**), differed significantly, ***t*(17.6) = 7.10, *p* < .001**.

The subjective vertical of the 10 students in philosophy (***Med* = 20.9 [19.3; 21]**), compared to that of the 10 students in psychology (***Med* = 20.9 [4.8; 8.9]**), differed significantly (***p* < .001, Mann-Whitney test**).

possibilité de mentionner le nom du test