

# Statistiques pour les sciences (MAT-4681)

Arthur Charpentier

# 11 - Tests d'hypothèse

été 2022

# Affirmations

- ▶ le candidat A sera réélu aux élections ce dimanche
- ▶ les femmes aiment autant regarder le hockey à la télévision que les hommes
- ▶ le médicament A est aussi efficace que le médicament B pour soigner les migraines
- ▶ 80% des gens qui prennent l'avion ont peur
- ▶ au moins 80% des gens qui prennent l'avion ont peur
- ▶ une digue de 2 mètres protège contre les crues centennaires

# Affirmation I

- ▶ le candidat A sera réélu aux élections ce dimanche

On peut interroger  $n$  personnes; pour  $i = 1, 2, \dots, n$ ,

$$y_i = \begin{cases} 1 & \text{si } i \text{ annonce qu'il votera pour A} \\ 0 & \text{si } i \text{ annonce qu'il ne votera pas pour A} \end{cases}$$

$Y = \mathbf{1}_A$  suit une loi de Bernoulli  $\mathcal{B}(p)$

A sera réélu si (et seulement si)  $p > 50\%$

on peut utiliser  $\hat{p} = \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$

on peut légitimement croire que A sera élu si  $\hat{p} \dots$  est grand

## Affirmation II

- ▶ les femmes aiment autant regarder le hockey à la télévision que les hommes

On peut interroger des femmes ( $m$ ) et des hommes ( $n$ )

$$x_i = \begin{cases} 1 & \text{si la femme } i \text{ aime regarder le hockey} \\ 0 & \text{si la femme } i \text{ n'aime pas regarder le hockey} \end{cases}$$

$$y_j = \begin{cases} 1 & \text{si l'homme } j \text{ aime regarder le hockey} \\ 0 & \text{si l'homme } j \text{ n'aime pas regarder le hockey} \end{cases}$$

$X = \mathbf{1}_{\text{NHL}}$  suit une loi de Bernoulli  $\mathcal{B}(p_x)$

$Y = \mathbf{1}_{\text{NHL}}$  suit une loi de Bernoulli  $\mathcal{B}(p_y)$

les femmes aiment autant regarder le hockey à la télévision que les hommes si (et seulement si)  $p_x \geq p_y$

$$\text{on peut utiliser } \hat{p}_x = \bar{x} = \frac{1}{m} \sum_{i=1}^m x_i \text{ et } \hat{p}_y = \bar{y} = \frac{1}{n} \sum_{j=1}^n y_j$$

## Affirmation III

- ▶ le médicament A est aussi efficace que le médicament B pour soigner les migraines

On peut interroger des personnes qui ont pris A ( $m$ ) et B ( $n$ )

$$x_i = \begin{cases} 1 & \text{si } i \text{ a pris A et a eu une migraine} \\ 0 & \text{si } i \text{ a pris A et n'a pas eu de migraine} \end{cases}$$

$$y_j = \begin{cases} 1 & \text{si } j \text{ a pris B et a eu une migraine} \\ 0 & \text{si } j \text{ a pris B et n'a pas eu de migraine} \end{cases}$$

$X = \mathbf{1}_A$  suit une loi de Bernoulli  $\mathcal{B}(p_x)$

$Y = \mathbf{1}_B$  suit une loi de Bernoulli  $\mathcal{B}(p_y)$

A est aussi efficace que B pour soigner les migraines si (et seulement si)  $p_x \geq p_y$

$$\text{on peut utiliser } \hat{p}_x = \bar{x} = \frac{1}{m} \sum_{i=1}^m x_i \text{ et } \hat{p}_y = \bar{y} = \frac{1}{n} \sum_{j=1}^n y_j$$

## Affirmation IV

- ▶ 80% des gens qui prennent l'avion ont peur

On peut interroger  $n$  personnes; pour  $i = 1, 2, \dots, n$ ,

$$y_i = \begin{cases} 1 & \text{si } i \text{ a peur en avion} \\ 0 & \text{si } i \text{ n'a pas peur en avion} \end{cases}$$

$Y = \mathbf{1}_A$  suit une loi de Bernoulli  $\mathcal{B}(p)$

80% des gens qui prennent l'avion ont peur si (et seulement si)  $p = 80\%$

$$\text{on peut utiliser } \hat{p} = \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

on peut légitimement croire que 80% des gens qui prennent l'avion ont peur si  $\hat{p} \dots$  est "proche" de 80%

# Affirmation V

- ▶ au moins 80% des gens qui prennent l'avion ont peur

On peut interroger  $n$  personnes; pour  $i = 1, 2, \dots, n$ ,

$$y_i = \begin{cases} 1 & \text{si } i \text{ a peur en avion} \\ 0 & \text{si } i \text{ n'a pas peur en avion} \end{cases}$$

$Y = \mathbf{1}_A$  suit une loi de Bernoulli  $\mathcal{B}(p)$

80% des gens qui prennent l'avion ont peur si (et seulement si)  $p \geq 80\%$

on peut utiliser  $\hat{p} = \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$

on peut légitimement croire que plus de 80% des gens qui prennent l'avion ont peur si  $\hat{p} \dots$  est plus grand que 80% ?

# Affirmation VI

- une digue de 2 mètres protège contre les crues centennaires

On peut observer le niveau annuel maximal d'un fleuve

$i = 1, 2, \dots, n$ ,

$y_i$  désigne le niveau de dépassement d'un fleuve

$Y$  a une loi  $F$ , inconnue. La digue de niveau  $s$  protège avec une probabilité  $1 - \alpha$  si  $\mathbb{P}[Y > s] = 1 - F(s) \leq 1 - \alpha$

La digue de niveau  $s$  protège contre les crues centennaires si  $F(s) \geq 99\%$

on peut utiliser  $\hat{F}(y) = \frac{1}{n} \sum_{i=1}^n y_i$

on peut légitimement croire que protège contre les crues centennaires si  $\hat{F}(2) \geq 99\%$ , non ?



# Inférence Ponctuelle & Intervalle de Confiance

XXXXX

# Hypothèse nulle & hypothèse alternative

- ▶ 80% des gens qui prennent l'avion ont peur

Cette affirmation sera l'hypothèse nulle, notée  $H_0$ .

Mais il convient de spécifier, si cette hypothèse n'est pas vérifiée, ce que pourrait être l'hypothèse alternative...

- ▶ 80% des gens qui prennent l'avion n'ont pas peur
- ▶ 75% des gens qui prennent l'avion ont peur
- ▶ il n'y a pas 80% des gens qui prennent l'avion qui ont peur
- ▶ moins de 80% des gens qui prennent l'avion ont peur

Cette affirmation sera l'hypothèse alternative, notée  $H_1$ .

On peut tenter de formaliser avec un modèle probabiliste.

Ici, on suppose que  $Y_i \sim F$  où  $F \in \mathcal{F} = \{F_\theta; \theta \in \Theta\}$ .

Bien entendu,  $\theta$  est inconnu, on dispose de  $\mathbf{y} = \{y_1, y_2, \dots, y_n\}$ .

# Hypothèse nulle & hypothèse alternative

- ▶ 80% des gens qui prennent l'avion ont peur  
 $H_0$ : hypothèse nulle,  $\theta = \theta_0$  (où  $\theta_0$  est donnée,  $\theta_0 = 80\%$ )
- ▶ 80% des gens qui prennent l'avion n'ont pas peur
- ▶ 75% des gens qui prennent l'avion ont peur  
 $H_1$ : hypothèse alternative,  $\theta = \theta_1$  (ex:  $\theta_1 = 20\%$  ou  $75\%$ )
- ▶ il n'y a pas 80% des gens qui prennent l'avion qui ont peur  
 $H_1$ : hypothèse alternative,  $\theta \neq \theta_0$  (ou  $\theta_1 \in \Theta_1 = \Theta \setminus \{\theta_0\}$ )
- ▶ moins de 80% des gens qui prennent l'avion ont peur  
 $H_1$ : hypothèse alternative,  $\theta < \theta_0$  ( $\theta_1 \in \Theta_1 = [-\infty; \theta_0)$ )

Plus généralement, avec  $\Theta_0 \cap \Theta_1 = \emptyset$ ,

- ▶  $H_0$ : hypothèse nulle,  $\theta \in \Theta_0$
- ▶  $H_1$ : hypothèse alternative,  $\theta \in \Theta_1$

# Les faiseurs de pluie

(histoire inspirée de Saporta (2006))

A partir de relevés obtenus après plusieurs décennies ont permis d'établir que dans une région, le niveau des pluies dans le Beauce, par ans, en mm, est  $y_i$ , où  $Y_i \sim \mathcal{N}(600, 100^2)$ .

Des entrepreneurs, appelés faiseurs de pluie, prétendaient pouvoir augmenter la pluviométrie, de l'ordre de 50 mm par an. Sur 9 années où l'expérience a été tenté, on a obtenu

$i$	2011	2012	2013	2014	2015	2016	2017	2018	2019
$y_i$	510	614	780	512	501	534	603	788	650

- ▶ La technique des faiseurs de pluie ne marche pas,  $H_0 : \mu = 600$
- ▶ La technique des faiseurs de pluie marche,  $H_1 : \mu = 650$

# Les faiseurs de pluie

Si  $H_0$  était vraie  $Y_i \sim \mathcal{N}(600, 100^2)$ ,

$$\overline{Y} = \frac{1}{9} \sum_{i=2011}^{2019} Y_i \sim \mathcal{N}\left(600, \frac{100^2}{9}\right)$$

Ici,  $\overline{Y}$  sera notre **statistique de test**. On aura (ici) tendance à **rejeter  $H_0$  au profit de  $H_1$**  si “ $\overline{y}$  est trop grand”.

On cherche donc un seuil  $k$  tel que  $\mathbb{P}$  tel que “trop grand” survienne avec 1 chance sur 20, i.e.

$$\mathbb{P}\left(\overline{Y} > k \mid \overline{Y} \sim \mathcal{N}\left(600, \frac{100^2}{9}\right)\right) = \frac{1}{20}$$

et

- ▶ Si  $\overline{y} > k$  on rejette  $H_0$ , et on retiendra  $H_1 : \mu = 650$  (avec 5% de chances de se tromper)
- ▶ Si  $\overline{y} < k$  on retient  $H_0$ ,  $H_0 : \mu = 600$

# Les faiseurs de pluie

Pour finir l'exercice, notons que, comme  $\alpha = 5\%$

$$k = 600 + u_{1-\alpha} \sqrt{\frac{100^2}{9}} = 600 + 1.64 \cdot \frac{100}{3} = 655$$

```
1 > 600+qnorm(.95)*100/3  
2 [1] 654.8285
```

Comme la règle de décision est

- ▶ si  $\bar{y} > 655$  on rejette  $H_0$ , et on retiendra  $H_1 : \mu = 650$  (avec 5% de chances de se tromper)
- ▶ si  $\bar{y} < 655$  on retient  $H_0$ ,  $H_0 : \mu = 600$

et que, numériquement,  $\bar{y} = 610.2$ , on retiendra  $H_0$ , et donc affirmer que les faiseurs ne pluie sont des charlatans.

Mais on peut se tromper....

# Les faiseurs de pluie et les erreurs

1. on peut se tromper en rejetant à tort  $H_0$ . Or par construction,

$$\mathbb{P}\left(\overline{Y} > 655 \mid \overline{Y} \sim \mathcal{N}\left(600, \frac{100^2}{9}\right)\right) = \mathbb{P}[\text{rejet } H_0 \mid H_0 \text{ vraie}] = 5\%$$

autrement dit, on a contrôlé cette erreur (dont la probabilité est noté  $\alpha$ ), dite **erreur de première espèce**.

2. on peut se tromper en acceptant à tort  $H_0$ , ce qui survient avec probabilité

$$\mathbb{P}[\text{accepter } H_0 \mid H_0 \text{ fausse}] = \mathbb{P}[\text{accepter } H_0 \mid H_1 \text{ vraie}]$$

$$= \mathbb{P}\left(\overline{Y} < 655 \mid \overline{Y} \sim \mathcal{N}\left(650, \frac{100^2}{9}\right)\right) = \Phi\left(\frac{655 - 650}{100/3}\right) = \Phi(0.15) \sim 56\%$$

cette probabilité est noté  $\beta$ , dite **erreur de seconde espèce**.

## Les faiseurs de pluie et les erreurs

On peut noter un lien avec les intervalles de confiance. Un intervalle de confiance bilatéral pour  $\mu$ , sur la base de  $\mathbf{y}$  serait

$$\left[ \bar{y} - u_{1-\alpha/2} \sqrt{\frac{100^2}{9}}; \bar{y} + u_{1-\alpha/2} \sqrt{\frac{100^2}{9}} \right] = [544.89; 675.55]$$

ou, comme la variance empirique de  $\mathbf{y}$  est 12438.69

$$\left[ \bar{y} - t_{8,1-\alpha/2} \sqrt{\frac{12439}{9}}; \bar{y} + t_{8,1-\alpha/2} \sqrt{\frac{12439}{9}} \right] = [524.49; 695.95]$$

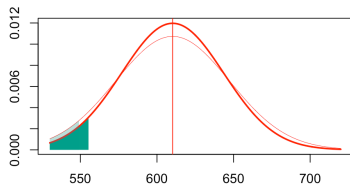
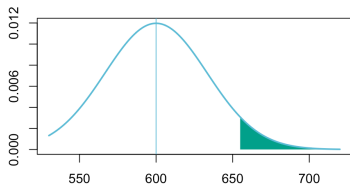
ou pour une version unilatérale de la forme  $[a, \infty)$ ,

$$\left[ \bar{y} - u_{1-\alpha} \sqrt{\frac{100^2}{9}}; \infty \right) = [555.39; \infty)$$

$$\text{ou } \left[ \bar{y} - t_{8,\alpha} \sqrt{\frac{12439}{9}}; \infty \right) = [541.09; \infty)$$



# Les faiseurs de pluie et les erreurs



Approche sur l'erreur

$$\mathbb{P}(\bar{Y} > k | H_0) = \frac{1}{20}$$

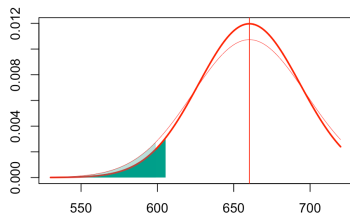
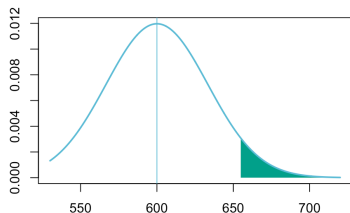
$$k = 600 + u_{1-\alpha} \sqrt{\frac{100^2}{9}} = 600 + 1.64 \cdot \frac{100}{3} = 655$$

Intervalle de confiance

$$\mathbb{P}\left(\mu \in \left[\bar{y} - u_{1-\alpha} \sqrt{\frac{100^2}{9}}; \infty\right) \mid H_1\right) = [555.39; \infty)$$

# Les faiseurs de pluie et les erreurs

mais si  $\bar{y} \rightarrow \bar{y} + 50$



Approche sur l'erreur

$$\mathbb{P}(\bar{Y} > k | H_0), k = 600 + u_{1-\alpha} \sqrt{\frac{100^2}{9}} = 655$$

Intervalle de confiance

$$\mathbb{P}\left(\mu \in \left[\bar{y} - u_{1-\alpha} \sqrt{\frac{100^2}{9}}; \infty\right) \middle| H_1\right) = [605.39; \infty)$$

## Les faiseurs de pluie et les erreurs

Dans le premier cas, on accepte que ce sont des faiseurs de pluie si

$$\bar{y} > 600 + u_{1-\alpha} \sqrt{\frac{100^2}{9}}$$

et dans le second cas on accepte que ce sont des faiseurs de pluie si

$$600 < \bar{y} - u_{1-\alpha} \sqrt{\frac{100^2}{9}}$$

... ce qui est équivalent.

On appellera l'intervalle

$$[k, \infty) = \left[ 600 + u_{1-\alpha} \sqrt{\frac{100^2}{9}}, \infty \right)$$

la **région critique**, notée  $W$ , ou **zone de rejet**.

# Hypothèses et décisions

On va maintenant formaliser ce qu'on vient de faire, en notant qu'ici

- ▶  $H_0$ : hypothèse nulle,  $\theta \in \Theta_0$  correspondait à  $\theta = \theta_0$
- ▶  $H_1$ : hypothèse alternative,  $\theta_1 \in \Theta_1$  correspondait à  $\theta = \theta_1$

On avait ici deux **hypothèses simples**. Le cas général  $\theta \in \Theta_0$  où  $\Theta_0 \neq \{\theta_0\}$  est appelé **hypothèse composite**. Les hypothèses compositives classiques sont (souvent)

$$\Theta_1 = \{\theta \in \Theta; \theta < \theta_0\}, \{\theta \in \Theta; \theta > \theta_0\} \text{ ou } \{\theta \in \Theta; \theta \neq \theta_0\}.$$

La décision sera aussi binaire, avec deux stratégies possibles

- ▶ rejeter l'hypothèse  $H_0$  (au profit de  $H_1$ )
- ▶ rejeter l'hypothèse  $H_1$  (au profit de  $H_0$ )

# Hypothèses et décisions

Pour prendre la décision on utilisera une **statistique de test**,  $T$ .

On va alors déterminer la forme de la **région critique**  $W$ , en fonction de  $H_1$ .

A partir de la probabilité  $\alpha$  (souvent 5%), on va déterminer les valeurs des bords de  $W$

On peut ensuite calculer la **probabilité d'erreur de seconde espèce**  $\beta$  (ou la puissance du test  $1 - \beta$ )

On calcule  $t = T(\mathbf{y})$  à partir des données, et **on prend une décision**

# Hypothèses et décisions

Sur l'histoire des faiseurs de pluie, la **statistique de test** utilisée est

$T(\mathbf{y}) = \bar{y}$ , on aurait pu prendre aussi  $T(\mathbf{y}) = \bar{y} - 600$  ou

$$T(\mathbf{y}) = \sqrt{9} \frac{\bar{y} - 600}{100}$$

Comme les hypothèses que l'on veut tester sont

- ▶ La technique des faiseurs de pluie ne marche pas,  $H_0 : \mu = 600$
- ▶ La technique des faiseurs de pluie marche,  $H_1 : \mu = 650$

la forme de la **région critique**  $W$  sera

$$W = \{t \in \mathbb{R}; t \text{ grand}\} = [k, \infty)$$

Heuristiquement, comme  $T$  est la moyenne

$$\left\{ \begin{array}{ll} H_0 : \theta = \theta_0 \text{ et } H_1 : \theta = \theta_1 > \theta_0 & : W = [k; \infty) \\ H_0 : \theta = \theta_0 \text{ et } H_1 : \theta > \theta_0 & : W = [k; \infty) \\ H_0 : \theta = \theta_0 \text{ et } H_1 : \theta = \theta_1 < \theta_0 & : W = (-\infty; k] \\ H_0 : \theta = \theta_0 \text{ et } H_1 : \theta = \theta_0 & : W = (-\infty; k] \\ H_0 : \theta = \theta_0 \text{ et } H_1 : \theta \neq \theta_0 & : W = (-\infty; k^-] \cup [k^+; \infty) \end{array} \right.$$

# Hypothèses et décisions

Dans le dernier cas

$$H_0 : \theta = \theta_0 \text{ et } H_1 : \theta \neq \theta_0 : W = (-\infty; k^-] \cup [k^+; \infty)$$

on peut aussi avoir une statistique  $T(\mathbf{y}) = |\bar{y} - \theta_0|$  ou  $(\bar{y} - \theta_0)^2$ ,  
et dans ce cas

$$H_0 : \theta = \theta_0 \text{ et } H_1 : \theta \neq \theta_0 : W = [k; \infty)$$

A partir de la probabilité  $\alpha$  (souvent 5%), on va déterminer les valeurs des bords de  $W$ , autrement dit  $k$  (ou  $k^-$  et  $k^+$  pour un test bilatéral). Pour nous, c'était

$$\mathbb{P}\left(\bar{Y} > k \mid \bar{Y} \sim \mathcal{N}\left(600, \frac{100^2}{9}\right)\right) = \alpha$$

mais plus généralement,

$$\mathbb{P}\left(T(\mathbf{Y}) > k \mid T(\mathbf{Y}) \sim G \text{ si } H_0 \text{ est vraie}\right) = 1 - G(k) = \alpha$$

soit  $k = G^{-1}(1 - \alpha)$ .

# Hypothèses et décisions

On peut ensuite calculer la **probabilité d'erreur de seconde espèce**  $\beta$  (ou la puissance du test  $1 - \beta$ ).

$$\beta = \mathbb{P}\left(\overline{Y} < k \mid \overline{Y} \sim \mathcal{N}\left(650, \frac{100^2}{9}\right)\right)$$

où  $k = 600 + u_{1-\alpha} \frac{100}{\sqrt{9}}$ , soit, avec  $\overline{Y} \sim \mathcal{N}\left(650, \frac{100^2}{9}\right)$ ,

$$\beta = \mathbb{P}\left(\underbrace{\sqrt{9} \frac{\overline{Y} - 650}{100}}_{Z \sim \mathcal{N}(0,1)} < \sqrt{9} \frac{k - 650}{100}\right) = \Phi\left(u_{1-\alpha} + \sqrt{9} \frac{600 - 650}{100}\right)$$

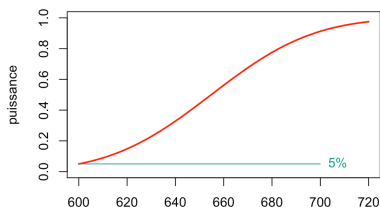
```
1 > pnorm(qnorm(.95)+sqrt(9)*(600-650)/100)
2 [1] 0.5575868
```



# Hypothèses et décisions

Plus généralement, on peut tracer la **puissance du test**, en fonction de  $\theta_1$  (pour  $\theta_1 > \theta_0$ ),

```
1 > puissance = fonction(t) 1-pnorm(qnorm(.95)+sqrt(9)  
    *(600-t)/100)
```



On calcule  $t = T(\mathbf{y})$  à partir des données, et **on prend une décision**.

Ici  $\bar{y} = 610.22$ , et comme  $k = 655$ , donc on ne rejette pas  $H_0$  car  $\bar{y} \notin W = [k, \infty)$ .

# Théorie de la décision

	$H_0$ est vraie	$H_1$ est vraie
rejeter l'hypothèse $H_1$	bonne décision	erreur de second type
rejeter l'hypothèse $H_0$	erreur de premier type	bonne décision

- $\alpha$  la probabilité d'erreur de première espèce,

$$\alpha = \mathbb{P}[T(\mathbf{Y}) \in W | H_0]$$

on parle aussi de seuil critique ou le seuil de signification

- $\beta$  la probabilité d'erreur de seconde espèce,

$$\beta = \mathbb{P}[T(\mathbf{Y}) \notin W | H_1]$$

et  $1 - \beta$  sera la puissance du test.

- en rejetant toujours  $H_0$ ,  $\alpha = 1$
- en rejetant toujours  $H_1$ ,  $\beta = 1$  (et la puissance est minimale)

# Erreurs de Type I et Type II

$$\hat{Y} = 0$$

NEGATIVE

$$\hat{Y} = 1$$

POSITIVE

$$Y = 0$$

NOT PREGNANT



$$Y = 1$$

PREGNANT



# $p$ -value ou probabilité critique

## Probabilité critique ( $p$ -value)

La probabilité critique ( $p$ -value) associée à une statistique de test est la probabilité d'observer des valeurs aussi ou plus extrêmes que la valeur observée dans l'échantillon sachant que  $H_0$  est vraie.

Elle est le plus petit seuil auquel on peut rejeter  $H_0$ . Pour faire simple, c'est le (plus petit) risque à encourir pour rejeter  $H_0$  et accepter  $H_1$ .

En fonction de celle-ci, la règle de décision peut se réécrire de la façon suivante : on rejete  $H_0$  (ou on accepte  $H_1$ ) si  $p\text{-valeur} < \alpha$ .

## $p$ -value ou probabilité critique

Ici, concrètement, on suppose  $H_0$  vérifiée

$$p = \mathbb{P}\left(\bar{Y} > \bar{y} \mid \bar{Y} \sim \mathcal{N}\left(600, \frac{100^2}{9}\right)\right) = \mathbb{P}\left(\underbrace{\sqrt{9} \frac{\bar{Y} - 600}{100}}_{Z \sim \mathcal{N}(0,1)} > \sqrt{9} \frac{\bar{y} - 600}{100}\right)$$

soit

```
1 > 1-pnorm(sqrt(9)*(mean(y)-600)/100)
2 [1] 0.3795486
```

Comme  $p = 37.95\% > \alpha = 5\%$ , on ne rejette pas  $H_0$ .

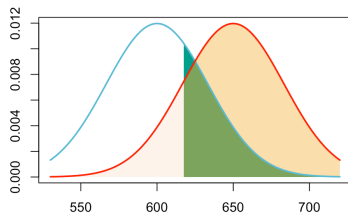
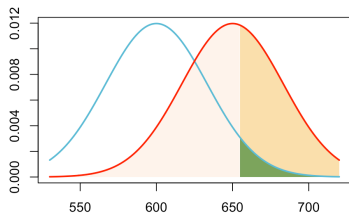
Si  $\bar{y} = 654.8$ ,  $p = 5\%$

```
1 > 1-pnorm(sqrt(9)*(654.8-600)/100)
2 [1] 0.050
```

autrement dit, si  $\bar{y} > 654.8$  (soit  $k$ ), on rejette  $H_0$ .

## $\alpha$ et $\beta$

En bleu, on a la distribution de  $\bar{Y}$  si  $H_0$  est vraie; et en rouge, on a la distribution de  $\bar{Y}$  si  $H_1$  est vraie



Si on veut réduire  $\beta$ , on le paye sur  $\alpha$ .

# Quelle statistique de test ?

Dans le cas des faiseurs de pluie, on a naturellement considéré  $T(\mathbf{y}) = \bar{y}$ . On peut aussi considérer **Neyman & Pearson (1933)**

## Neyman-Pearson

Pour un test de la forme  $H_0 : \theta = \theta_0$  contre  $H_1 : \theta = \theta_1$ , on considérera

$$T = \frac{\mathcal{L}(\theta_0; \mathbf{y})}{\mathcal{L}(\theta_1; \mathbf{y})}$$

et la région de rejet sera  $W = (-\infty, \gamma]$ .

Ici on a un modèle Gaussien (par hypothèse),  $\mathcal{L}(\theta; \mathbf{y})$  vaut

$$\prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_i - \theta)^2}{\sigma^2}\right) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{\sigma^2} \sum_{i=1}^n (y_i - \theta)^2\right)$$

## Quelle statistique de test ?

$$T = \frac{\mathcal{L}(\theta_0; \mathbf{y})}{\mathcal{L}(\theta_1; \mathbf{y})} = \exp \left( -\frac{1}{\sigma^2} \sum_{i=1}^n (y_i - \theta_0)^2 + \frac{1}{\sigma^2} \sum_{i=1}^n (y_i - \theta_1)^2 \right)$$

soit

$$T = \exp \left( \frac{1}{\sigma^2} \left[ \sum_{i=1}^n (y_i - \theta_1)^2 - \sum_{i=1}^n (y_i - \theta_0)^2 \right] \right)$$

$$T = \exp \left( \frac{1}{\sigma^2} \left[ \sum_{i=1}^n 2y_i(\theta_0 - \theta_1) + (n\theta_0^2 - n\theta_1^2) \right] \right)$$

Aussi,  $T < \gamma$  signifie  $\log(T) < \log(\gamma)$ ,

$$\sum_{i=1}^n 2y_i(\theta_0 - \theta_1) < \sigma^2 \log(k) - n(\theta_1^2 - \theta_0^2)$$

$$2(\theta_0 - \theta_1)\bar{y} < \frac{\sigma^2}{n} \log(k) - (\theta_1^2 - \theta_0^2)$$



## Quelle statistique de test ?

$$2(\theta_0 - \theta_1)\bar{y} < \frac{\sigma^2}{n} \log(k) - (\theta_1^2 - \theta_0^2)$$

donc,

- ▶ si  $\theta_1 < \theta_0$ ,  $T < \gamma$  signifie  $\bar{y} < k$
- ▶ si  $\theta_1 > \theta_0$ ,  $T < \gamma$  signifie  $\bar{y} > k$

où  $k$  est de la forme

$$\frac{\sigma^2 \log(k)}{2n|\theta_1 - \theta_0|} + (\theta_1 + \theta_0)$$

Dans le cas d'un modèle Gaussien, un test sur la moyenne à la Neyman-Pearson revient à regarder la valeur de  $\bar{y}$ .

## Quelques tests

Modèle Gaussien avec variance connue ( $\sigma^2$ )

Test  $H_0 : \mu = \mu_0$  contre  $H_1 : \mu = \mu_1, \mathcal{N}(\mu, \sigma^2)$

Soit  $\mathbf{x} = \{x_1, \dots, x_n\}$  de loi  $\mathcal{N}(\mu, \sigma^2)$ .

Pour tester  $H_0 : \mu = \mu_0$  contre  $H_1 : \mu = \mu_1$ , on utilise

$$Z = \sqrt{n} \frac{\bar{x} - \mu_0}{\sigma}$$

- ▶ si  $\mu_1 > \mu_0$ , on rejette  $H_0$  si  $z > \Phi^{-1}(1 - \alpha) = u_{1-\alpha}$
- ▶ si  $\mu_1 < \mu_0$ , on rejette  $H_0$  si  $z < \Phi^{-1}(\alpha) = u_\alpha$

Pour illustrer, simulons un échantillon  $\mathcal{N}(0, 1)$

```
1 > set.seed(1)
2 > x = rnorm(30)
3 > mean(x)
4 [1] 0.08245817
```

# Quelques tests

La statistique de test est  $z$

```
1 > (z = sqrt(30)*(mean(x)-0)/1)
2 [1] 0.451642
```

Si on teste  $H_0 : \mu = 0$  contre  $H_0 : \mu = \mu_1 < 0$ , la  $p$ -value est

```
1 > pnorm(z)
2 [1] 0.6742365
```

et  $W = (-\infty; -1.64]$  si  $\alpha = 5\%$

```
1 > qnorm(c(0, .05))
2 [1] -Inf -1.644854
```

- ▶ Comme  $z \notin W$ , on ne rejette pas  $H_0$
- ▶ Comme  $p > \alpha$ , on ne rejette pas  $H_0$   
(avec un niveau de confiance  $\alpha = 5\%$ )

# Quelques tests

La statistique de test est  $z$

```
1 > (z = sqrt(30)*(mean(x)-0)/1)
2 [1] 0.451642
```

Si on teste  $H_0 : \mu = 0$  contre  $H_0 : \mu = \mu_1 > 0$ , la  $p$ -value est

```
1 > 1-pnorm(z)
2 [1] 0.3257635
```

et  $W = [1.64; +\infty)$  si  $\alpha = 5\%$

```
1 > qnorm(c(.95,1))
2 [1] 1.644854      Inf
```

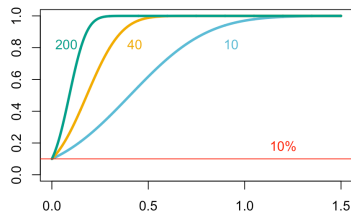
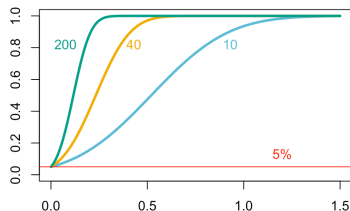
- ▶ Comme  $z \notin W$ , on ne rejette pas  $H_0$
- ▶ Comme  $p > \alpha$ , on ne rejette pas  $H_0$   
(avec un niveau de confiance  $\alpha = 5\%$ )

# Quelques tests

Pour le test unilatéral à droite ( $H_1 : \mu = \mu_1 > \mu_0$ ), on peut montrer

$$\text{puissance} = 1 - \beta = 1 - \Phi \left( u_{1-\alpha} - \frac{\mu_1 - \mu_0}{\sigma} \sqrt{n} \right)$$

donc si on trace la puissance en fonction de l'écart normalisé  $\frac{\mu_1 - \mu_0}{\sigma}$ , quand  $\alpha = 5\%$  ou  $10\%$ ,



## Quelques tests

Modèle Gaussien avec variance connue ( $\sigma^2$ )

Test  $H_0 : \mu = \mu_0$  contre  $H_1 : \mu \neq \mu_0$ ,  $\mathcal{N}(\mu, \sigma^2)$

Soit  $\mathbf{x} = \{x_1, \dots, x_n\}$  de loi  $\mathcal{N}(\mu, \sigma^2)$ .

Pour tester  $H_0 : \mu = \mu_0$  contre  $H_1 : \mu \neq \mu_0$ , on utilise

$$Z = \sqrt{n} \frac{\bar{x} - \mu_0}{\sigma}$$

► on rejette  $H_0$  si  $|z| > \Phi^{-1}(1 - \alpha/2) = u_{1-\alpha/2}$

Dans ce cas, on peut montrer que, si  $\Delta = \mu - \mu_0$ ,

$$\beta = \Phi\left(u_{1-\alpha/2} - \frac{\Delta}{\sigma}\sqrt{n}\right) - \Phi\left(u_{\alpha/2} - \frac{\Delta}{\sigma}\sqrt{n}\right)$$

## Quelques tests

On utilise le même échantillon. La statistique de test reste

```
1 > (z = sqrt(30)*(mean(x)-0)/1)
2 [1] 0.451642
```

Si on teste  $H_0 : \mu = 0$  contre  $H_0 : \mu \neq 0$ , la  $p$ -value est

$$p = \mathbb{P}[|Z| > |z|] = 2 \cdot (\mathbb{P}[Z > |z|]) = 2 \cdot (1 - \mathbb{P}[Z \leq |z|]) = 2 \cdot (1 - \Phi(|z|))$$

```
1 > 2*(1-pnorm(abs(z)))
2 [1] 0.6515269
```

et  $W = (-\infty; -1.96] \cup [1.96; +\infty)$  si  $\alpha = 5\%$ , i.e.

```
1 > qnorm(c(0, .025, .975, 1))
2 [1]          -Inf -1.959964  1.959964          Inf
```

► Comme  $z \notin W$ , on ne rejette pas  $H_0$

► Comme  $p > \alpha$ , on ne rejette pas  $H_0$   
(avec un niveau de confiance  $\alpha = 5\%$ )

## Quelques tests

On peut simuler quelques échantillons suivant des lois  $\mathcal{N}(0, 1)$

```
1 > simu = function(i){
2 +   set.seed(i)
3 +   x = rnorm(30)
4 +   z = sqrt(30)*(mean(x)-0) /1
5 +   p = c(mean(x),z,pnorm(z),1-pnorm(z),2*(1-pnorm(
      abs(z))))
6 +   names(p) = c("moyenne","z","<0",">0","<>0")
7 +   p
8 + }
9 > simu(1)
10   moyenne          z          <0          >0          <>0
11 0.08245817 0.45164200 0.67423655 0.32576345 0.65152691
```

Parfois, on va rejeter à tort  $H_0$  (car on simule vraiment des échantillons de moyenne 0)

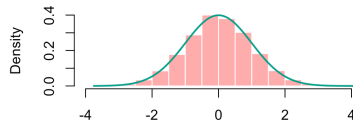
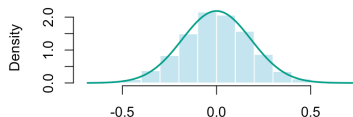
```
1 > simu(4)
2   moyenne          z          <0          >0          <>0
3 0.48718726 2.66843455 0.99618972 0.00381028 0.00762056
```



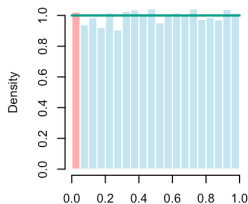
# Quelques tests

On peut simuler beaucoup d'échantillons suivant des lois  $\mathcal{N}(0, 1)$

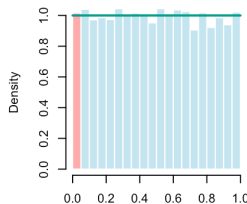
```
1 > S = Vectorize(simu)(1:1e4)
2 > hist(S[1,])
3 > hist(S[2,])
```



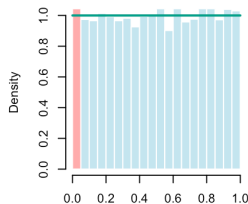
```
1 > hist(S[3,])
2 > hist(S[4,])
3 > hist(S[5,])
```



p-value,  $\mu < 0$



p-value,  $\mu > 0$



p-value,  $\mu \neq 0$

## Quelques tests

Modèle Gaussien avec variance connue ( $\sigma^2$ )

Test  $H_0 : \mu_x - \mu_y = \mu_0$  contre  $H_1 : \mu_x - \mu_y = \mu_1, \mathcal{N}(\mu., \sigma^2)$

Soient  $\mathbf{x} = \{x_1, \dots, x_m\}$  de loi  $\mathcal{N}(\mu_x, \sigma_x^2)$  et  $\mathbf{y} = \{y_1, \dots, y_n\}$  de loi  $\mathcal{N}(\mu_y, \sigma_y^2)$ .

Pour tester  $H_0 : \mu_x - \mu_y = \mu_0$  contre  $H_1 : \mu_x - \mu_y = \mu_1$ , on utilise

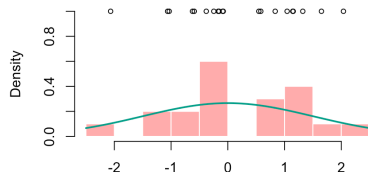
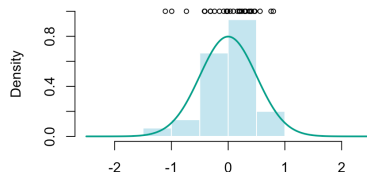
$$Z = \frac{(\bar{x} - \bar{y}) - \mu_0}{\sqrt{\frac{\sigma_x^2}{m} + \frac{\sigma_y^2}{n}}}$$

- ▶ si  $\mu_1 > \mu_0$ , on rejette  $H_0$  si  $z > \Phi^{-1}(1 - \alpha) = u_{1-\alpha}$
- ▶ si  $\mu_1 < \mu_0$ , on rejette  $H_0$  si  $z < \Phi^{-1}(\alpha) = u_\alpha$

# Quelques tests

On simule ici deux échantillons,  $\mathcal{N}(0, 0.5^2)$  et  $\mathcal{N}(0, 1.5^2)$

```
1 > set.seed(1)
2 > x = rnorm(30, 0, .5)
3 > mean(x)
4 [1] 0.04122909
5 > y = rnorm(20, 0, 1.5)
6 > mean(y)
7 [1] 0.1911502
```



## Quelques tests

La statistique de test est  $z$

```
1 > (z = (mean(x)-mean(y))/sqrt(.5^2/30+1.5^2/20))  
2 [1] -0.4312899
```

Si on teste  $H_0 : \mu_x - \mu_y = 0$  contre  $H_0 : \mu_x - \mu_y = \mu_1 > 0$ , la  $p$ -value est

```
1 > 1-pnorm(z)  
2 [1] 0.6668712
```

et  $W = [1.64; +\infty)$  si  $\alpha = 5\%$

```
1 > qnorm(c(.95,1))  
2 [1] 1.644854      Inf
```

- ▶ Comme  $z \notin W$ , on ne rejette pas  $H_0$
- ▶ Comme  $p > \alpha$ , on ne rejette pas  $H_0$   
(avec un niveau de confiance  $\alpha = 5\%$ )

## Quelques tests

Si on teste  $H_0 : \mu_x - \mu_y = 0$  contre  $H_0 : \mu_x - \mu_y = \mu_1 < 0$ , la  $p$ -value est

```
1 > pnorm(z)
2 [1] 0.3331288
```

et  $W = (-\infty, -1.64]$  si  $\alpha = 5\%$

```
1 > qnorm(c(0, .05))
2 [1] -Inf -1.644854
```

- ▶ Comme  $z \notin W$ , on ne rejette pas  $H_0$
- ▶ Comme  $p > \alpha$ , on ne rejette pas  $H_0$   
(avec un niveau de confiance  $\alpha = 5\%$ )

## Quelques tests

Modèle Gaussien avec variance connue ( $\sigma^2$ )

Test  $H_0 : \mu_x - \mu_y = \mu_0$  contre  $H_1 : \mu_x - \mu_y \neq \mu_0$ ,  $\mathcal{N}(\mu., \sigma^2)$

Soient  $\mathbf{x} = \{x_1, \dots, x_m\}$  de loi  $\mathcal{N}(\mu_x, \sigma_x^2)$  et  $\mathbf{y} = \{y_1, \dots, y_n\}$  de loi  $\mathcal{N}(\mu_y, \sigma_y^2)$ .

Pour tester  $H_0 : \mu_x - \mu_y = \mu_0$  contre  $H_1 : \mu_x - \mu_y \neq \mu_0$ , on utilise

$$Z = \frac{(\bar{x} - \bar{y}) - \mu_0}{\sqrt{\frac{\sigma_x^2}{m} + \frac{\sigma_y^2}{n}}}$$

► on rejette  $H_0$  si  $|z| > \Phi^{-1}(1 - \alpha/2) = u_{1-\alpha/2}$

## Quelques tests

La statistique de test est (toujours)  $z$

```
1 > (z = (mean(x)-mean(y))/sqrt(.5^2/30+1.5^2/20))  
2 [1] -0.4312899
```

Si on teste  $H_0 : \mu_x - \mu_y = 0$  contre  $H_0 : \mu_x - \mu_y \neq 0$ , la  $p$ -value est

```
1 > 2*(1-pnorm(abs(z)))  
2 [1] 0.6662576
```

et  $W = (-\infty; -1.96] \cup [1.96; +\infty)$  si  $\alpha = 5\%$ , i.e.

```
1 > qnorm(c(0,.025,.975,1))  
2 [1] -Inf -1.959964 1.959964 Inf
```

▶ Comme  $z \notin W$ , on ne rejette pas  $H_0$

▶ Comme  $p > \alpha$ , on ne rejette pas  $H_0$   
(avec un niveau de confiance  $\alpha = 5\%$ )

## Quelques tests

Mais dans la vraie vie,  $\sigma$  est rarement connue...

Modèle Gaussien avec variance inconnue

Test  $H_0 : \mu = \mu_0$  contre  $H_1 : \mu = \mu_1, \mathcal{N}(\mu, \sigma^2)$

Soit  $\mathbf{x} = \{x_1, \dots, x_n\}$  de loi  $\mathcal{N}(\mu, \sigma^2)$ .

Pour tester  $H_0 : \mu = \mu_0$  contre  $H_1 : \mu = \mu_1$ , on utilise

$$T = \sqrt{n} \frac{\bar{x} - \mu_0}{s}, \quad s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

- ▶ si  $\mu_1 > \mu_0$ , on rejette  $H_0$  si  $t > T_{n-1}^{-1}(1 - \alpha)$
- ▶ si  $\mu_1 < \mu_0$ , on rejette  $H_0$  si  $t < T_{n-1}^{-1}(\alpha)$

où  $T_\nu$  est la fonction de répartition de la loi de Student  $\text{Std}(\nu)$ .



## Quelques tests

```
1 > set.seed(1)
2 > x = rnorm(30)
3 > mean(x)
4 [1] 0.08245817
```

La statistique de test est  $t$

```
1 > (t = sqrt(30)*(mean(x)-0)/sd(x))
2 [1] 0.4887261
```

Si on teste  $H_0 : \mu = 0$  contre  $H_0 : \mu = \mu_1 < 0$ , la  $p$ -value est

```
1 > pt(t, df = 30-1)
2 [1] 0.6856444
```

et  $W = (-\infty; -1.7]$  si  $\alpha = 5\%$

```
1 > qt(c(0, .05), df = 30-1)
2 [1] -Inf -1.699127
```

- ▶ Comme  $z \notin W$ , on ne rejette pas  $H_0$
- ▶ Comme  $p > \alpha$ , on ne rejette pas  $H_0$   
(avec un niveau de confiance  $\alpha = 5\%$ )

## Quelques tests

On peut aussi calculer l'intervalle de confiance (asymétrique) pour  $\mu$ , l'espérance de  $Y$

$$\left( -\infty; \bar{x} + T_{n-1}^{-1}(1 - \alpha) \sqrt{\frac{s^2}{n}} \right], \quad s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

```
1 > mean(x) + qt(.95, df=29) * sd(x) / sqrt(30)
2 [1] 0.3691359
```

$\mu_0 = 0$  est dans cet intervalle  $(-\infty, 0.37]$ ,

- Comme  $0 \in IC_{\alpha}$ , on ne rejette pas  $H_0$   
(avec un niveau de confiance  $\alpha = 5\%$ )

## Quelques tests

On peut aussi utiliser `t.test`,

```
1 > t.test(x, mu=0, alternative = "less")
2
3   One Sample t-test
4
5 data:  y
6 t = 0.48873, df = 29, p-value = 0.6856
7 alternative hypothesis: true mean is less than 0
8 95 percent confidence interval:
9    -Inf 0.3691359
```

L'option `alternative = "less"` signifie  $\mu_1 < \mu_0$ .

## Quelques tests

L'option alternative = "greater" signifie  $\mu_1 > \mu_0$ .

```
1 > t.test(x, mu=0, alternative = "greater")
2
3   One Sample t-test
4
5 data:  x
6 t = 0.48873, df = 29, p-value = 0.3144
7 alternative hypothesis: true mean is greater than 0
8 95 percent confidence interval:
9  -0.2042196      Inf
10 sample estimates:
11  mean of x
12 0.08245817
```

- ▶ Comme  $p > \alpha$ , on ne rejette pas  $H_0$
- ▶ Comme  $0 \in IC_\alpha$ , on ne rejette pas  $H_0$   
(avec un niveau de confiance  $\alpha = 5\%$ )

# Quelques tests

## Modèle Gaussien avec variance inconnue

Test  $H_0 : \mu = \mu_0$  contre  $H_1 : \mu \neq \mu_0$ ,  $\mathcal{N}(\mu, \sigma^2)$

Soit  $\mathbf{x} = \{x_1, \dots, x_n\}$  de loi  $\mathcal{N}(\mu, \sigma^2)$ .

Pour tester  $H_0 : \mu = \mu_0$  contre  $H_1 : \mu \neq \mu_0$ , on utilise

$$T = \sqrt{n} \frac{\bar{x} - \mu_0}{s}, \quad s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

► on rejette  $H_0$  si  $|t| > T_{n-1}^{-1}(1 - \alpha/2)$

où  $T_\nu$  est la fonction de répartition de la loi de Student  $\text{Std}(\nu)$ .

## Quelques tests

Là encore, on peut utiliser la fonction `t.test`

```
1 > t.test(x, mu=0, alternative = "two.sided")
2
3   One Sample t-test
4
5 data:  y
6 t = 0.48873, df = 29, p-value = 0.6287
7 alternative hypothesis: true mean is not equal to 0
8 95 percent confidence interval:
9  -0.2626142  0.4275306
```

Notons au passage qu'on a un intervalle de confiance (bilatéral) pour  $\mu$ ,

$$\left[ \bar{x} - T_{n-1}^{-1}(1 - \alpha/2) \sqrt{\frac{s^2}{n}}; \bar{x} + T_{n-1}^{-1}(1 - \alpha/2) \sqrt{\frac{s^2}{n}} \right]$$

```
1 > qt(.975, df=29)
2 [1] 2.04523
```

## Quelques tests

On peut utiliser des vraies données.

La pression artérielle se mesure à l'aide de deux chiffres : Le premier nombre, appelé **pression artérielle systolique**, mesure la pression dans vos artères lorsque votre cœur bat. Le deuxième chiffre, appelé tension diastolique, mesure la pression dans vos artères lorsque votre cœur se repose entre deux battements.

```
1 > loc = "http://freakonometrics.free.fr/MAT4681/  
   blood_pressure.txt"  
2 > download.file(loc, "blood_pressure.txt")  
3 > blood_pressure = read.table("blood_pressure.txt",  
   header=TRUE, sep=",")  
4 > mean(blood_pressure$mmhg)  
5 [1] 130
```

# Quelques tests

```
1 > t.test(blood_pressure$mmhg, mu=140, alternative="
    less")
2
3     One Sample t-test
4
5 data:  blood_pressure$mmhg
6 t = -3.8693, df = 54, p-value = 0.0001481
7 alternative hypothesis: true mean is less than 140
8 95 percent confidence interval:
9     -Inf 134.3253
10 sample estimates:
11 mean of x
12     130
```



## Quelques tests

Modèle Gaussien avec variances inconnues (mais égales)

Test  $H_0 : \mu_x - \mu_y = \mu_0$  contre  $H_1 : \mu_x - \mu_y = \mu_1$ ,  $\mathcal{N}(\mu., \sigma^2)$

Soient  $\mathbf{x} = \{x_1, \dots, x_m\}$  de loi  $\mathcal{N}(\mu_x, \sigma_x^2)$  et  $\mathbf{y} = \{y_1, \dots, y_n\}$  de loi  $\mathcal{N}(\mu_y, \sigma_y^2)$ .

Pour tester  $H_0 : \mu_x - \mu_y = \mu_0$  contre  $H_1 : \mu_x - \mu_y = \mu_1$ , on utilise

$$T = \frac{(\bar{x} - \bar{y}) - \mu_0}{s \sqrt{\frac{1}{m} + \frac{1}{n}}}, \quad s^2 = \frac{(m-1)s_x^2 + (n-1)s_y^2}{m+n-2}$$

- ▶ si  $\mu_1 > \mu_0$ , on rejette  $H_0$  si  $z > T_{m+n-2}^{-1}(1 - \alpha)$
- ▶ si  $\mu_1 < \mu_0$ , on rejette  $H_0$  si  $z < T_{m+n-2}^{-1}(\alpha)$

## Quelques tests

Là encore, on peut utiliser la fonction `t.test`.

Si on veut tester  $H_0 : \mu_x - \mu_y = 0$  contre  $H_1 : \mu_x - \mu_y < 0$

```
1 > t.test(x, y, mu=0, alternative = "less")
2
3   Welch Two Sample t-test
4
5 data:  x and y
6 t = -0.37905, df = 37.551, p-value = 0.3534
7 alternative hypothesis: true difference in means is
   less than 0
8 95 percent confidence interval:
9    -Inf 0.3749006
10 sample estimates:
11  mean of x  mean of y
12 0.08245817 0.19115017
```

$H_1 : \mu_x - \mu_y < 0$  se traduit par  
true difference in means is less than 0

## Quelques tests

Là encore, on peut utiliser la fonction `t.test`.

On peut aussi tester  $H_0 : \mu_x - \mu_y = 0$  contre  $H_1 : \mu_x - \mu_y > 0$

```
1 > t.test(x, y, mu=0, alternative = "greater")
2
3 Welch Two Sample t-test
4
5 data: x and y
6 t = -0.37905, df = 37.551, p-value = 0.6466
7 alternative hypothesis: true difference in means is
   greater than 0
8 95 percent confidence interval:
9  -0.5922846      Inf
10 sample estimates:
11  mean of x mean of y
12 0.08245817 0.19115017
```

$H_1 : \mu_x - \mu_y > 0$  se traduit par  
true difference in means is greater than 0

## Quelques tests

On note que quelle que soit l'hypothèse alternative  $H_1$  ( $H_1 : \mu_x - \mu_y < 0$  ou  $H_1 : \mu_x - \mu_y > 0$ ),

- ▶ Comme  $p > \alpha$ , on ne rejette pas  $H_0$
- ▶ Comme  $0 \in IC_\alpha$ , on ne rejette pas  $H_0$   
(avec un niveau de confiance  $\alpha = 5\%$ )

Si les deux variances  $\sigma_x^2$  et  $\sigma_y^2$  sont inconnues (soyons réaliste), mais qu'on peut supposer égales on peut proposer un autre test.

## Quelques tests

Modèle Gaussien avec variances inconnues (mais égales)

Test  $H_0 : \mu_x - \mu_y = \mu_0$  contre  $H_1 : \mu_x - \mu_y \neq \mu_0$ ,  $\mathcal{N}(\mu., \sigma^2)$

Soient  $\mathbf{x} = \{x_1, \dots, x_m\}$  de loi  $\mathcal{N}(\mu_x, \sigma^2)$  et  $\mathbf{y} = \{y_1, \dots, y_n\}$  de loi  $\mathcal{N}(\mu_y, \sigma^2)$  (avec la même variance  $\sigma^2$ ).

Pour tester  $H_0 : \mu_x - \mu_y = \mu_0$  contre  $H_1 : \mu_x - \mu_y \neq \mu_0$ , on utilise

$$T = \frac{(\bar{x} - \bar{y}) - \mu_0}{s \sqrt{\frac{1}{m} + \frac{1}{n}}}, \quad s^2 = \frac{(m-1)s_x^2 + (n-1)s_y^2}{m+n-2}$$

► on rejette  $H_0$  si  $|z| > T_{m+n-2}^{-1}(1 - \alpha/2) = u_{1-\alpha/2}$

## Quelques tests

Dans la fonction `t.test`, il est possible d'utiliser l'option `var.equal = TRUE`

```
1 > set.seed(1)
2 > x = rnorm(30,0,1)
3 > y = rnorm(20,0,1)
4 > t.test(x, y, mu=0, alternative = "two.sided", var.
    equal = TRUE)
5
6 Two Sample t-test
7
8 data: x and y
9 t = -0.18554, df = 48, p-value = 0.8536
10 alternative hypothesis: true difference in means is
    not equal to 0
11 95 percent confidence interval:
12 -0.5323591 0.4424086
13 sample estimates:
14 mean of x mean of y
15 0.08245817 0.12743344
```

# Quelques tests

## Sur les données de pression artérielle

```
1 > t.test(blood_pressure$mmhg ~ blood_pressure$status ,  
2         mu=0, alternative="two.sided", var.equal=TRUE)  
3  
4     Two Sample t-test  
5  
6 data:  blood_pressure$mmhg by blood_pressure$status  
7 t = -10.468, df = 53, p-value = 1.66e-14  
8 alternative hypothesis: true difference in means  
9    between group 0 and group 1 is not equal to 0  
10 95 percent confidence interval:  
11 -37.31328 -25.31339  
12  
13 sample estimates:  
14 mean in group 0 mean in group 1  
15      112.9200      144.2333
```

# Quelques tests

```
1 > t.test(blood_pressure$mmhg ~ blood_pressure$status)
2
3 Welch Two Sample t-test
4
5 data: blood_pressure$mmhg by blood_pressure$status
6 t = -10.451, df = 50.886, p-value = 2.887e-14
7 alternative hypothesis: true difference in means
   between group 0 and group 1 is not equal to 0
8 95 percent confidence interval:
9  -37.32904 -25.29763
10 sample estimates:
11 mean in group 0 mean in group 1
12      112.9200      144.2333
```



## Quelques tests

```
1 > t.test(blood_pressure$mmhg ~ blood_pressure$status,  
2         mu=0, alternative="two.sided", var.equal=FALSE)  
3  
4 Welch Two Sample t-test  
5  
6 data: blood_pressure$mmhg by blood_pressure$status  
7 t = -10.451, df = 50.886, p-value = 2.887e-14  
8 alternative hypothesis: true difference in means  
9 between group 0 and group 1 is not equal to 0  
10 95 percent confidence interval:  
11 -37.32904 -25.29763  
12 sample estimates:  
13 mean in group 0 mean in group 1  
14 112.9200 144.2333
```

# Quelques tests

## Modèle Gaussien

Test  $H_0 : \sigma^2 = \sigma_0^2$  contre  $H_1 : \sigma^2 = \sigma_1^2, \mathcal{N}(\mu, \sigma^2)$

Soit  $\mathbf{x} = \{x_1, \dots, x_n\}$  de loi  $\mathcal{N}(\mu, \sigma^2)$ .

Pour tester  $H_0 : \sigma^2 = \sigma_0^2$  contre  $H_1 : \sigma^2 = \sigma_1^2$ , on utilise

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}, \quad s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

► si  $\sigma_1^2 > \sigma_0^2$ , on rejette  $H_0$  si  $\chi^2 > Q_{n-1}^{-1}(1 - \alpha)$

► si  $\sigma_1^2 < \sigma_0^2$  on rejette  $H_0$  si  $\chi^2 < Q_{n-1}^{-1}(\alpha)$

où  $Q_\nu$  est la fonction de répartition de la loi du chi-deux,  $\chi^2(\nu)$ .

## Quelques tests

$H_0 : \sigma^2 = \sigma_0^2$  peut se tester avec la fonction `varTest` de `library(EnvStats)`.  $H_0 : \sigma^2 = 1$  contre  $H_1 : \sigma^2 > 1$

```
1 > EnvStats::varTest(x, alternative="greater", conf.  
  level = 0.95, sigma.squared = 1)  
2 $statistic  
3 Chi-Squared  
4    24.76598  
5  
6 $p.value  
7 [1] 0.6903324  
8  
9 $estimate  
10 variance  
11 0.8539993  
12  
13 $conf.int  
14      LCL      UCL  
15 0.5819489      Inf  
16 attr(,"conf.level")
```

## Quelques tests

$H_0 : \sigma^2 = \sigma_0^2$  peut se tester avec la fonction `varTest` de `library(EnvStats)`.  $H_0 : \sigma^2 = 1$  contre  $H_1 : \sigma^2 < 1$

```
1 > EnvStats::varTest(x, alternative="less", conf.level
  = 0.95, sigma.squared = 1)
2 $statistic
3 Chi-Squared
4     24.76598
5
6 $p.value
7 [1] 0.3096676
8
9 $estimate
10 variance
11 0.8539993
12
13 $conf.int
14      LCL      UCL
15 0.000000 1.398547
```

# Quelques tests

## Modèle Gaussien

Test  $H_0 : \sigma^2 = \sigma_0^2$  contre  $H_1 : \sigma^2 \neq \sigma_0^2$ ,  $\mathcal{N}(\mu, \sigma^2)$

Soit  $\mathbf{x} = \{x_1, \dots, x_n\}$  de loi  $\mathcal{N}(\mu, \sigma^2)$ .

Pour tester  $H_0 : \sigma^2 = \sigma_0^2$  contre  $H_1 : \sigma^2 \neq \sigma_0^2$ , on utilise

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}, \quad s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

► on rejette  $H_0$  si  $\chi^2 < Q_{n-1}^{-1}(\alpha/2)$  ou  $\chi^2 > Q_{n-1}^{-1}(1 - \alpha/2)$

où  $Q_\nu$  est la fonction de répartition de la loi du chi-deux,  $\chi^2(\nu)$ .

## Quelques tests

$H_0 : \sigma^2 = \sigma_0^2$  peut se tester avec la fonction `varTest` de `library(EnvStats)`.  $H_0 : \sigma^2 = 1$  contre  $H_1 : \sigma^2 \neq 1$

```
1 > EnvStats::varTest(x, alternative="two.sided", conf.  
  level = 0.95, sigma.squared = 1)  
2 $statistic  
3 Chi-Squared  
4    24.76598  
5  
6 $p.value  
7 [1] 0.6193352  
8  
9 $estimate  
10 variance  
11 0.8539993  
12  
13 $conf.int  
14      LCL      UCL  
15 0.541661 1.543333
```

# Quelques tests

## Modèle Gaussien

Test  $H_0 : \sigma_x^2 = \sigma_y^2$  contre  $H_1 : \sigma_x^2 \neq \sigma_y^2$ ,  $\mathcal{N}(\mu, \sigma^2)$

Soient  $\mathbf{x} = \{x_1, \dots, x_m\}$  de loi  $\mathcal{N}(\mu_x, \sigma_x^2)$  et  $\mathbf{y} = \{y_1, \dots, y_n\}$  de loi  $\mathcal{N}(\mu_y, \sigma_y^2)$ .

Pour tester  $H_0 : \sigma_x^2 = \sigma_y^2$  contre  $H_1 : \sigma_x^2 \neq \sigma_y^2$ , on utilise

$$F = \frac{s_x^2}{s_y^2}, \quad s_x^2 = \frac{1}{m-1} \sum_{i=1}^m (x_i - \bar{x})^2, \quad s_y^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2$$

► on rejette  $H_0$  si  $f < F_{m-1, n-1}^{-1}(\alpha/2)$  ou  
 $f > F_{m-1, n-1}^{-1}(1 - \alpha/2)$

où  $F_{\nu_1, \nu_2}$  est la fonction de répartition de la loi de Fisher,  $\mathcal{F}(\nu_1, \nu_2)$ .

## Quelques tests

$H_0 : \sigma_x^2 = \sigma_y^2$  signifie que ratio = 1 dans var.test,

```
1 > set.seed(1)
2 > x = rnorm(30,0,1)
3 > y = rnorm(20,0,1)
4 > var.test(x, y, ratio = 1, alternative = "two.sided")
5
6   F test to compare two variances
7
8 data:  x and y
9 F = 1.7871, num df = 29, denom df = 19, p-value =
   0.1896
10 alternative hypothesis: true ratio of variances is not
   equal to 1
11 95 percent confidence interval:
12  0.7440377 3.9875896
13 sample estimates:
14 ratio of variances
15      1.787136
```



# Quelques tests

## Sur les données de pression artérielle

```
1 > var.test(blood_pressure$mmhg ~ blood_pressure$status
2             ,alternative="two.sided")
3
4     F test to compare two variances
5
6 data:  blood_pressure$mmhg by blood_pressure$status
7 F = 1.0363, num df = 24, denom df = 29, p-value =
8     0.918
9
10 alternative hypothesis: true ratio of variances is not
11     equal to 1
12 95 percent confidence interval:
13     0.4811235 2.2980313
14
15 sample estimates:
16 ratio of variances
17     1.036343
```

## Quelques tests

Avec deux jeux de données **indépendants**, on utilise le fait que la somme (et la différence) de loi normales indépendantes suit une loi normales, et la somme de lois du chi-deux indépendantes suit un loi du chi-deux.

Si les données sont appariées, il est possible que les observations ne soient pas indépendantes entre les deux groupes.

```
1 > loc= "http://freakonometrics.free.fr/MAT4681/iq.txt"
2 > download.file(loc,"iq.txt")
3 > iq = read.table("iq.txt", header=TRUE, sep=",")
4 > cor(iq$IQ1,iq$IQ2)
5 [1] 0.9948492
```

$$\text{Notons } r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \sum_{i=1}^n (y_i - \bar{y})^2}} \text{ la corrélation empirique.}$$

## Quelques tests

Test  $H_0 : \mu_x - \mu_y = \mu_0$  contre  $H_1 : \mu_x - \mu_y = \mu_1, \mathcal{N}(\mu, \sigma^2)$

Soient  $\{(x_1, y_1), \dots, (x_n, y_n)\}$  tels que  $x$  et  $y$  soient de loi  $\mathcal{N}(\mu_x, \sigma_x^2)$  et  $\mathcal{N}(\mu_y, \sigma_y^2)$ .

Pour tester  $H_0 :: \mu_x - \mu_y = \mu_0$  contre  $H_1 :: \mu_x - \mu_y = \mu_1$ , on utilise

$$Z = \sqrt{n} \frac{(\bar{x} - \bar{y}) - \mu_0}{\sqrt{s_x^2 + s_y^2 - 2\text{cov}(x, y)}}$$

Si  $H_0 :: \mu_x - \mu_y = \mu_0$  est vraie,  $Z$  suit une loi  $\mathcal{N}(0, 1)$ .

- ▶ si  $\mu_1 > \mu_0$  on rejette  $H_0$  si  $z > \Phi^{-1}(1 - \alpha)$
- ▶ si  $\mu_1 < \mu_0$  on rejette  $H_0$  si  $z < \Phi^{-1}(\alpha)$

## Quelques tests

Sur les données de QI,  $\text{Var}(x) \sim \text{Var}(y) + 10$

```
1 > t.test(iq$IQ1,iq$IQ2,mu=-10,alternative="two.sided",
2         paired=TRUE)
3     Paired t-test
4
5 data:  iq$IQ1 and iq$IQ2
6 t = -1.2854, df = 19, p-value = 0.2141
7 alternative hypothesis: true difference in means is
8     not equal to -10
9 95 percent confidence interval:
10  -11.051338  -9.748662
11 sample estimates:
12      mean of the differences
13      -10.4
```

## Quelques tests

Test  $H_0 : \sigma_x^2 = \sigma_y^2$  contre  $H_1 : \sigma_x^2 \neq \sigma_y^2$ ,  $\mathcal{N}(\mu, \sigma^2)$

Soient  $\{(x_1, y_1), \dots, (x_n, y_n)\}$  tels que  $x$  et  $y$  soient de loi  $\mathcal{N}(\mu_x, \sigma_x^2)$  et  $\mathcal{N}(\mu_y, \sigma_y^2)$ .

Pour tester  $H_0 : \sigma_x^2 = \sigma_y^2$  contre  $H_1 : \sigma_x^2 \neq \sigma_y^2$ , on utilise

$$T = \frac{\sqrt{n-2}(s_x^2 - s_y^2)}{\sqrt{4(1-r^2)s_x^2 s_y^2}}$$

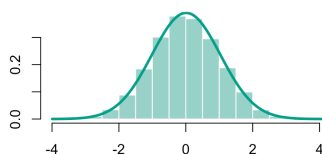
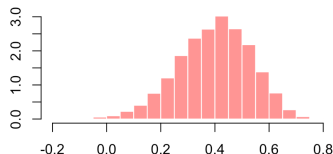
Si  $H_0 : \sigma_x^2 = \sigma_y^2$  est vraie,  $T$  suit une loi  $Std(n-2)$ .

► on rejette  $H_0$  si  $|t| > T_{n-2}^{-1}(1 - \alpha/2)$

# Quelques tests

On peut faire des simulations pour vérifier

```
1 > r = 0.4
2 > z =matrix(NA,10000,2)
3 > for(s in 1:10000){
4 + x = rnorm(n=40)
5 + y = r*x+sqrt(1-r^2)*rnorm(n=40)
6 + cr = cor(x,y)
7 + t = sqrt(40-2)*(var(x)-var(y))/sqrt(4*(1-cr^2)*var(x)
8   )*var(y))
9 + z[s,] = c(cr,t)
9 +}
```



# Quelques tests

Ici,

```
1 > std1 = sd(iq$IQ1)
2 > std2 = sd(iq$IQ2)
3 > n = length(iq$IQ1)
4 > r = cor(iq$IQ1,iq$IQ2)
5 > (t = (sqrt(df)*(std1^2-std2^2))/(4*(1-r^2)*std1^2*
      std2^2))
6 [1] 0.007821987
7 > 2*pt(-abs(t),n-2)
8 [1] 0.9937999
```

# Test d'ajustement de loi

## PP Plot, $X \sim F$

Soit  $\{x_1, \dots, x_n\}$  un échantillon, notons  $\{x_{(1)}, \dots, x_{(n)}\}$  la version ordonnée. Le PP-plot est le nuage de points

$$\left\{ \left( \frac{i}{n+1}, F(x_{(i)}) \right), i = 1, \dots, n \right\}$$

## Test $H_0 : X \sim F$ contre $H_1 : X \not\sim F$ , PP Plot

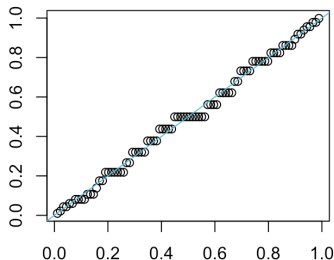
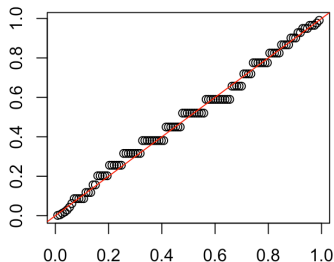
Soit  $\{x_1, \dots, x_n\}$  un échantillon, notons  $\{x_{(1)}, \dots, x_{(n)}\}$  la version ordonnée. Si  $H_0 : X \sim F$  est vraie, les points sont alignés sur la première bissectrice ( $y = x$ ).



# Test d'ajustement de loi

On peut tester si  $X$  suit une loi normale en faisant un PP plot avec  $F$  la fonction de répartition de la loi normale  $\mathcal{N}(\bar{x}, \hat{s}^2)$

```
1 > x = Davis$height[Davis$sex == "F"]  
2 > u = (1:length(x))/(1+length(x))  
3 > plot(u, pnorm(sort(x), mean(x), sd(x)))  
4 > abline(a = 0, b = 1)  
5 > y = Davis$height[Davis$sex == "M"]
```



# Test d'ajustement de loi

## QQ Plot, $X \sim F$

Soit  $\{x_1, \dots, x_n\}$  un échantillon, notons  $\{x_{(1)}, \dots, x_{(n)}\}$  la version ordonnée. Le QQ-plot est le nuage de points

$$\left\{ \left( F^{-1} \left( \frac{i}{n+1} \right), x_{(i)} \right), i = 1, \dots, n \right\}$$

Test  $H_0 : X \sim F$  contre  $H_1 : X \not\sim F$ , QQ Plot

Soit  $\{x_1, \dots, x_n\}$  un échantillon, notons  $\{x_{(1)}, \dots, x_{(n)}\}$  la version ordonnée. Si  $H_0 : X \sim F$  est vraie, les points sont alignés sur la première bissectrice ( $y = x$ ).

# Test d'ajustement de loi

## QQ Plot Gaussien

Soit  $\{x_1, \dots, x_n\}$  un échantillon, notons  $\{x_{(1)}, \dots, x_{(n)}\}$  la version ordonnée. Le QQ-plot Gaussien est le nuage de points

$$\left\{ \left( \Phi^{-1} \left( \frac{i}{n+1} \right), x_{(i)} \right), i = 1, \dots, n \right\}$$

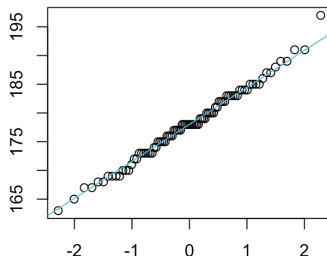
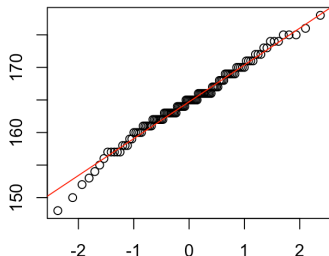
Test  $H_0 : X \sim \mathcal{N}$  contre  $H_1 : X \not\sim \mathcal{N}$ , QQ Plot Gaussien

Soit  $\{x_1, \dots, x_n\}$  un échantillon, notons  $\{x_{(1)}, \dots, x_{(n)}\}$  la version ordonnée. Si  $H_0 : X \sim \mathcal{N}(\mu, \sigma^2)$  est vraie, les points sont alignés sur la droite  $y = \mu + \sigma x$ .

# Test d'ajustement de loi

On peut utiliser la fonction `qqnorm` dans le cas Gaussien

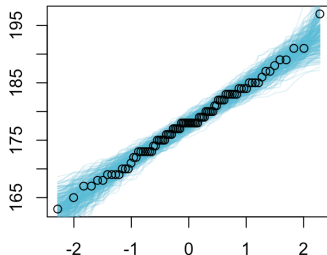
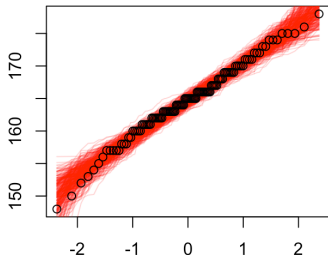
```
1 > x = Davis$height[Davis$sex == "F"]
2 > qqnorm(x)
3 > u = (1:length(x))/(1+length(x))
4 > plot(qnorm(u), sort(x))
5 > abline(a = mean(x), b = sd(x))
6 > y = Davis$height[Davis$sex == "M"]
```



# Test d'ajustement de loi

On peut simuler des échantillons Gaussien (même moyenne, même variance) pour avoir des bornes de confiances

```
1 > for(s in 1:500){  
2 +   xs=rnorm(length(x), mean(x), sd(x))  
3 +   lines(qnorm(u), sort(xs))  
4 + }
```



Dans le cas continu (par exemple Gaussien) on pourra utiliser le test de Kolmogorov Smirnov.

# Test d'ajustement de loi

Test  $H_0 : X \sim F$  contre  $H_1 : X \not\sim F$ , Kolmogorov-Smirnov

Soit  $\mathbf{x} = \{x_1, \dots, x_n\}$  de loi  $F_X$ .

Pour tester  $H_0 : F_X = F$  contre  $H_1 : F_X \neq F$ , on utilise

$$D = \sup_{t \in \mathbb{R}} \{|\hat{F}_X(t) - F(t)|\}$$

► on rejette  $H_0$  si  $d > K_n^{-1}(1 - \alpha)$

où  $K_\nu$  est la loi du test de Komogorov-Smirnov.

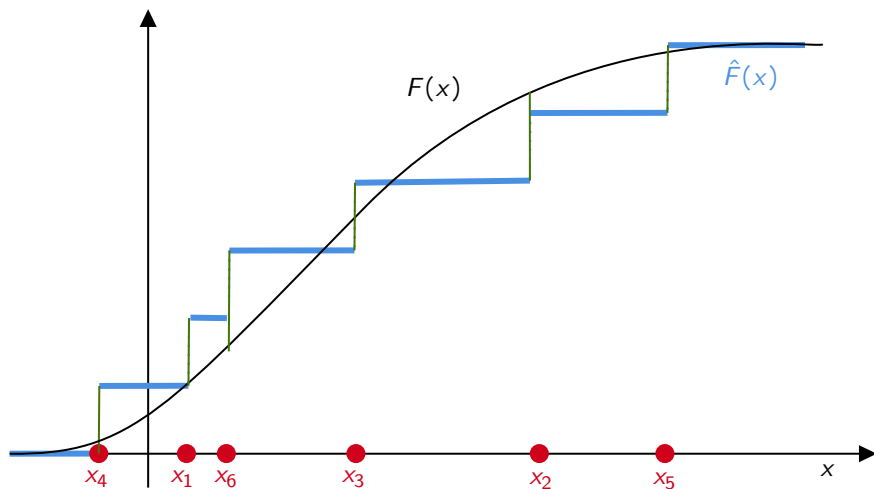
Statistique Kolmogorov-Smirnov

$$D = \max_{1 \leq i \leq n} \left\{ F(x_{(i)}) - \frac{i-1}{n}, \frac{i}{n} - F(x_{(i)}) \right\}$$

où  $\{x_{(1)}, \dots, x_{(n)}\}$  est l'échantillon ordonné.

# Test d'ajustement de loi

$$D = \max_{1 \leq i \leq n} \left\{ F(x_{(i)}) - \frac{i-1}{n}, \frac{i}{n} - F(x_{(i)}) \right\}$$



# Test d'ajustement de loi

Numériquement, on utilise `ks.test` sous R

```
1 > ks.test(x, "pnorm", mean = 165, sd =5)
2
3   One-sample Kolmogorov-Smirnov test
4
5 data:  x
6 D = 0.081455, p-value = 0.4472
7 alternative hypothesis: two-sided
```

Même si on ne connaît pas la loi suivie par  $D$  (sous  $H_0$ ), il suffit de regarder la  $p$ -value.

Ici  $p = 44.7\%$  donc on ne rejette pas  $H_0 : X \sim \mathcal{N}(165, 5^2)$ .

**Note** si on veut tester  $H_0 : X \sim \mathcal{N}(\bar{x}, s^2)$ , la  $p$ -value n'est pas juste est doit être modifiée. On parle alors de test de Lilliefors



## Test d'ajustement d'une famille de loi ★★★

Pour le test de  $H_0 : X \sim \mathcal{N}(\bar{x}, s^2)$ ,

```
1 > ks.test(x,"pnorm",mean=mean(x),sd=sd(x))
2
3 One-sample Kolmogorov-Smirnov test
4
5 data: x
6 D = 0.070851, p-value = 0.6275
7 alternative hypothesis: two-sided
```

```
1 > library(nortest)
2 > lillie.test(x)
3
4 Lilliefors (Kolmogorov-Smirnov) normality test
5
6 data: x
7 D = 0.070851, p-value = 0.1818
```

Sinon, il existe des dizaines de tests dits de normalité (test d'Anderson-Darling, test d'Agostino-Pearson, test de Geary, test de Shapiro-Wilk ou de Shapiro-Francia)

# Test d'ajustement d'une famille de loi ★★★

```
1 > ad.test(x)
2
3 Anderson-Darling normality test
4
5 data:  x
6 A = 0.38576, p-value = 0.3858
7 > sf.test(x)
8
9 Shapiro-Francia normality test
10
11 data:  x
12 W = 0.98901, p-value = 0.4258
```

Ces deux tests ne permettent pas de rejeter  $H_0$  (qui est ici une hypothèse de normalité de la taille des filles).

# Test d'ajustement de loi

Test  $H_0 : F_X = F_Y$  contre  $H_1 : F_X \neq F_Y$ , KS

Soient  $\mathbf{x} = \{x_1, \dots, x_m\}$  de loi  $F_X$  et  $\mathbf{y} = \{y_1, \dots, y_n\}$  de loi  $F_Y$ .

Pour tester  $H_0 : F_X = F_Y$  contre  $H_1 : F_X \neq F_Y$ , on utilise

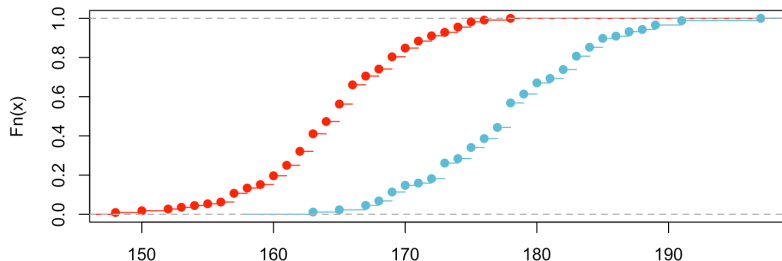
$$D = \sup_{t \in \mathbb{R}} \{|\hat{F}_X(t) - \hat{F}_Y(t)|\}$$

► on rejette  $H_0$  si  $d > K_{m,n}^{-1}(1 - \alpha)$

où  $K_{\nu_1, \nu_2}$  est la loi du test de Komogorov-Smirnov.

# Test d'ajustement de loi

```
1 > ks.test(x,y)
2
3 Two-sample Kolmogorov-Smirnov test
4
5 data: x and y
6 D = 0.7289, p-value < 2.2e-16
7 alternative hypothesis: two-sided
8 > plot(ecdf(x), xlim = range(c(x,y)))
9 > plot(ecdf(y), add = TRUE)
```



# Test d'ajustement de loi ★★ ★

Test  $H_0 : F_X = F_Y$  contre  $H_1 : F_X \neq F_Y$ , KS,  $m$  et  $n$  grands

Soient  $\mathbf{x} = \{x_1, \dots, x_m\}$  de loi  $F_X$  et  $\mathbf{y} = \{y_1, \dots, y_n\}$  de loi  $F_Y$ , avec  $m$  et  $n$  grands

Pour tester  $H_0 : F_X = F_Y$  contre  $H_1 : F_X \neq F_Y$ , on utilise

$$D = \sup_{t \in \mathbb{R}} \{|\hat{F}_X(t) - \hat{F}_Y(t)|\}$$

► on rejette  $H_0$  si  $d > c(\alpha) \sqrt{\frac{n+m}{n \cdot m}}$

où  $c(10\%) = 1.224$ ,  $c(5\%) = 1.358$  et  $c(1\%) = 1.628$ .

# The “ $p < 5\%$ ” Dogma

“if  $p$  is between 10% and 90% there is certainly no reason to suspect the hypothesis tested. If it is below 2% it is strongly indicated that the hypothesis fails to account for the whole of facts [...] We shall not often be astray if we draw a conventional line at 5%”

Ronald Fisher

see [La guerre des étoiles](#),  $p$ -value and statistical practice or [It's time to talk about ditching statistical significance](#)

<u>P-VALUE</u>	<u>INTERPRETATION</u>
0.001	HIGHLY SIGNIFICANT
0.01	
0.02	
0.03	
0.04	SIGNIFICANT
0.049	
0.050	OH CRAP. REDO CALCULATIONS.
0.051	ON THE EDGE OF SIGNIFICANCE
0.06	
0.07	HIGHLY SUGGESTIVE, SIGNIFICANT AT THE $P < 0.10$ LEVEL
0.08	
0.09	
0.099	HEY, LOOK AT THIS INTERESTING SUBGROUP ANALYSIS
$\geq 0.1$	