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 centenaires

Affirmation I

- le candidat A sera réélu aux élections ce dimanche

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

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

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

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

on peut utiliser $\widehat{p} = \overline{y} = \dfrac{1}{n}\sum_{i=1}^{n} y_i$

on peut légitimement croire que A sera élu si $\widehat{p}$... 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) $p_x \geq p_y$

on peut utiliser $\widehat{p}_x = \overline{x} = \dfrac{1}{m}\sum_{i=1}^{m} x_i$ et $\widehat{p}_y = \overline{x} = \dfrac{1}{n}\sum_{j=1}^{n} y_i$

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)

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

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

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

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

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

on peut utiliser $\widehat{p}_x = \overline{x} = \dfrac{1}{m}\displaystyle\sum_{i=1}^{m} x_i$ et $\widehat{p}_y = \overline{x} = \dfrac{1}{n}\displaystyle\sum_{j=1}^{n} y_j$

Affirmation IV

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

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

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

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

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

on peut utiliser $\widehat{p} = \overline{y} = \dfrac{1}{n}\sum_{i=1}^{n} y_i$

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

Affirmation V

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

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

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

$Y = \mathbf{1}_A$ 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 $\widehat{p} = \overline{y} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} y_i$

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

Affirmation VI

- une digue de 2 mètres protège contre les crues centenaires

On peut observer le niveau annuel maximal d'un fleuve $i = 1, 2, \cdots, n$,

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 centenaires si $F(s) \geq 99\%$

on peut utiliser $\widehat{F}(y) = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} y_i$

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

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, θ est inconnu, on dispose de $\boldsymbol{y} = \{y_1, y_2, \cdots, y_n\}$.

Hypothèse nulle & hypothèse alternative

- 80% des gens qui prennent l'avion ont peur
  H_0: hypothèse nulle, $\theta = \theta_0$ (où θ_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 $\theta_1 \in \Theta_1 = \Theta \backslash \{\theta_0\}$)
- 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, $\theta_1 \in \Theta_1$

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.

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

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 $\alpha = 5\%$

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

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

Comme la règle de décision est

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

et que, numériquement, $\overline{y} = 610.2$, 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,

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

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

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

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

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

cette probabilité est noté β, 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 μ, sur la base de $\boldsymbol{y}$ serait

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

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

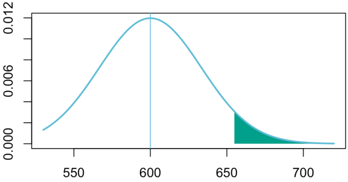
$$\left[\overline{y}-t_{8,1-\alpha/2}\sqrt{\frac{12439}{9}};\overline{y}+t_{8,1-\alpha/2}\sqrt{\frac{12439}{9}}\right]=\left[524.49;695.95\right]$$

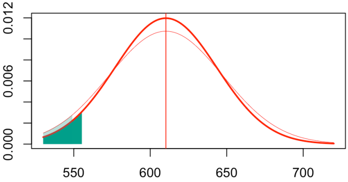
ou pour une version une unilatérale de la forme $[a,\infty)$,

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

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

Les faiseurs de pluie et les erreurs





Approche sur l'erreur

$$\mathbb{P}\left(\overline{Y} > k \middle| H_0\right) = \frac{1}{20}$$

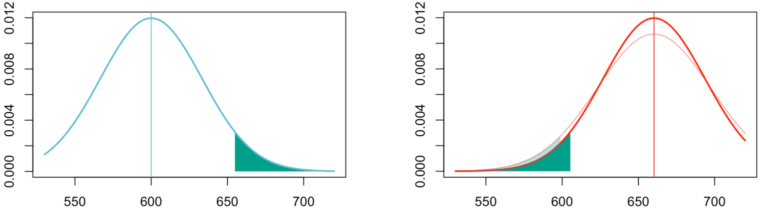
$$k = 600 + u_{1-\alpha}\sqrt{\frac{100^2}{9}} = 600 + 1.64 \cdot \frac{100}{3} = 655$$

Intervalle de confiance

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

Les faiseurs de pluie et les erreurs

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



Approche sur l'erreur

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

Intervalle de confiance

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

Les faiseurs de pluie et les erreurs

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

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

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

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

... 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 α (souvent 5%), on va déterminer les valeurs des bords de W

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

On calcule $t = T(\boldsymbol{y})$ 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 $T(\boldsymbol{y}) = \overline{y}$, on aurait pu prendre aussi $T(\boldsymbol{y}) = \overline{y} - 600$ ou $T(\boldsymbol{y}) = \sqrt{9}\dfrac{\overline{y} - 600}{100}$

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

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

Heuristiquement, comme T est la moyenne

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

Hypothèses et décisions

Dans le dernier cas

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

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

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

A partir de la probabilité α (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

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

mais plus généralement,

$$\mathbb{P}\Big(T(\boldsymbol{Y}) > k \Big| T(\boldsymbol{Y}) \sim G \text{ si } H_0 \text{ est vraie}\Big) = 1 - G(k) = \alpha$$

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

Hypothèses et décisions

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

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

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

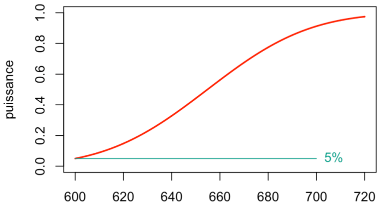
$$\beta = \mathbb{P}\Big(\underbrace{\sqrt{9}\frac{\overline{Y}-650}{100}}_{Z\sim\mathcal{N}(0,1)} < \sqrt{9}\frac{k-650}{100}\Big) = \Phi\left(u_{1-\alpha} + \sqrt{9}\frac{600-650}{100}\right)$$

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

Hypothèses et décisions

Plus généralement, on peut tracer la puissance du test, en fonction de θ_1 (pour $\theta_1 > \theta_0$),

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



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

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

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

- α la probabilité d'erreur de première espèce,

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

  on parle aussi de seuil critique ou le seuil de signification

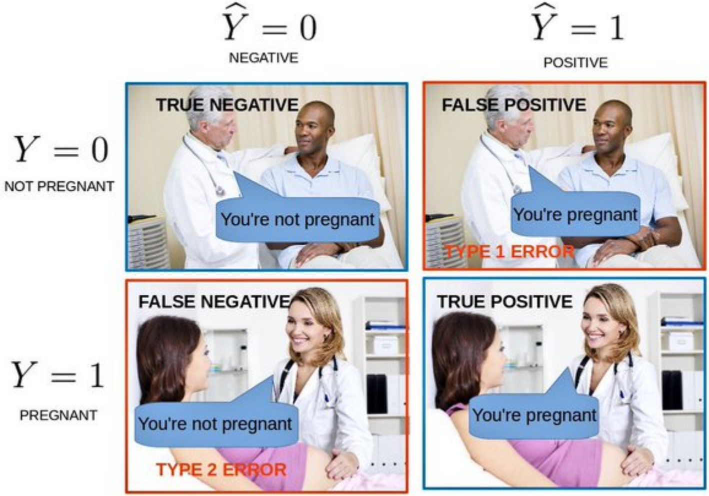
- β la probabilité d'erreur de seconde espèce,

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

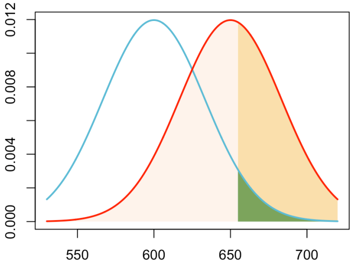
  et $1 - \beta$ sera la puissance du test.

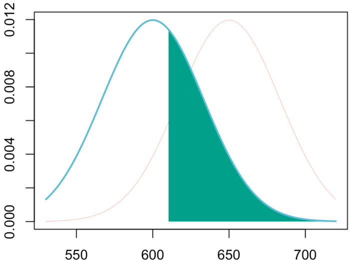
- en rejetant toujours H_0, $\alpha = 0$
- en rejetant toujours H_1, $\beta = 0$ (et la puissance est maximale)

Erreurs de Type I et Type II



p-value ou région de rejet (de H_0)





Pour l'instant, on a les distributions de $\overline{Y}$ sous H_0 ($\mathcal{N}(600, \star)$) et sous H_1 ($\mathcal{N}(650, \star)$)

- en se fixant α, on détermine k, borne de la région de rejet $W = [k, \infty) : \mathbb{P}[\overline{Y} \in W | H_0] = \mathbb{P}[\overline{Y} > k | H_0] = \alpha$

mais on peut aussi calculer la probabilité d'observer des valeurs aussi ou plus extrêmes que la valeur observée dans l'échantillon

- $\mathbb{P}[\overline{Y} > \overline{y} | H_0] = p$: si $p > \alpha$, $\overline{Y} \notin W$

p-value ou probabilité critique

Probabilité critique (p-value)

La probabilité critique (p-value) associéee 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-valeur$< \alpha$.

p-value ou probabilité critique

Ici, concrètement, on suppose H_0 vérifiée

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

soit

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

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

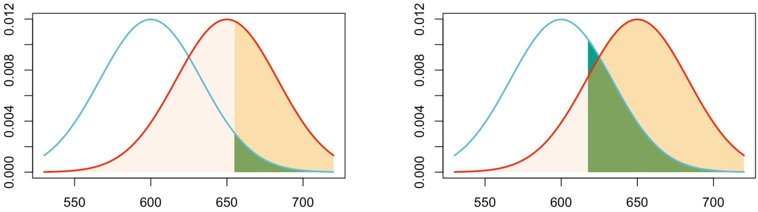
Si $\overline{y} = 654.8$, $p = 5\%$

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

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

α et β

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



Si on veut réduire β, on le paye sur α.

Quelle statistique de test ?

Dans le cas des faiseurs de pluie, on a naturellement considéré $T(\boldsymbol{y}) = \overline{y}$. 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

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

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

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

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

soit

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

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

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

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

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

Quelle statistique de test ?

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

donc,

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

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 $\overline{y}$.

Quelques tests

Modèle Gaussien avec variance connue (σ^2)

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

Soit $\boldsymbol{x} = \{x_1, \cdots, x_n\}$ de loi $\mathcal{N}(\mu, \sigma^2)$.
Pour tester $H_0 : \mu = \mu_0$ contre $H_1 : \mu = \mu_1$, on utilise

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

- 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 $\mathcal{N}(0, 1)$

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

Quelques tests

La statistique de test est z

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

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

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

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

```
> qnorm(c(0,.05))
[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

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

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

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

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

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

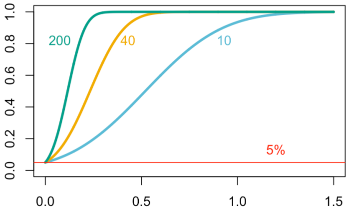
- Comme $z \notin W$, on ne rejette pas H_0
- Comme $p > \alpha$, on ne rejette pas H_0
  (avec un niveau de confiance $\alpha = 5\%$

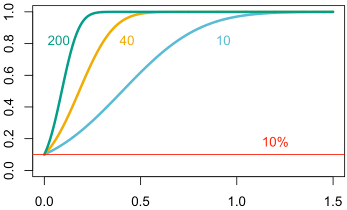
Quelques tests

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 $\dfrac{\mu_1 - \mu_0}{\sigma}$, quand $\alpha = 5\%$ ou 10%,





Quelques tests

Modèle Gaussien avec variance connue (σ^2)

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

Soit $\boldsymbol{x} = \{x_1, \cdots, x_n\}$ de loi $\mathcal{N}(\mu, \sigma^2)$.
Pour tester $H_0 : \mu = \mu_0$ contre $H_1 : \mu \neq \mu_0$, on utilise

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

- 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

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

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

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

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

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

```
> qnorm(c(0,.025,.975,1))
[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 $\mathcal{N}(0,1)$

```
> simu = function(i){
+      set.seed(i)
+      x = rnorm(30)
+      z = sqrt(30)*(mean(x)-0) /1
+      p = c(mean(x),z,pnorm(z),1-pnorm(z),2*(1-pnorm(
     abs(z))))
+      names(p) = c("moyenne","z","<0",">0","<>0")
+      p
+ }
> simu(1)
   moyenne          z         <0         >0        <>0
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)

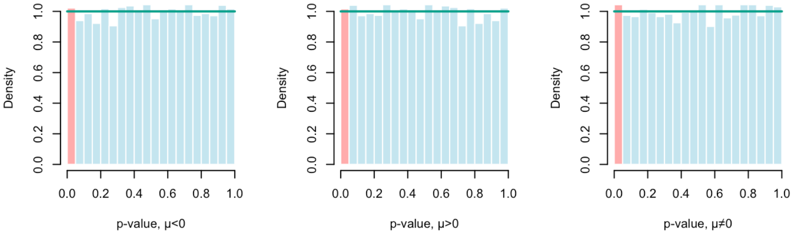
```
> simu(4)
   moyenne          z         <0         >0        <>0
0.48718726 2.66843455 0.99618972 0.00381028 0.00762056
```

Quelques tests

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

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


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



Quelques tests

Modèle Gaussien avec variance connue ($\sigma_\cdot^2$)

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

Soient $\boldsymbol{x} = \{x_1, \cdots, x_m\}$ de loi $\mathcal{N}(\mu_x, \sigma_x^2)$ et $\boldsymbol{y} = \{y_1, \cdots, y_n\}$ 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

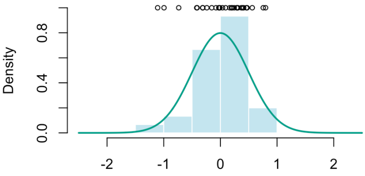
$$Z = \frac{(\overline{x} - \overline{y}) - \mu_0}{\sqrt{\dfrac{\sigma_x^2}{m} + \dfrac{\sigma_y^2}{n}}}$$

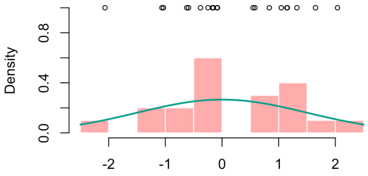
- si $\mu_1 > \mu_0$, on rejette H_0 si $z > \Phi^{-1}(1-\alpha) = u_{1-\alpha}$
- si $\mu_1 < \mu_0$, on rejette H_0 si $z < \Phi^{-1}(\alpha) = u_\alpha$

Quelques tests

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

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





Quelques tests

La statistique de test est z

```
> (z = (mean(x)-mean(y))/sqrt(.5^2/30+1.5^2/20))
[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-pnorm(z)
[1] 0.6668712
```

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

```
> qnorm(c(.95,1))
[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

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

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

```
> qnorm(c(0,.05))
[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 ($\sigma_\cdot^2$)

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

Soient $\boldsymbol{x} = \{x_1, \cdots, x_m\}$ de loi $\mathcal{N}(\mu_x, \sigma_x^2)$ et $\boldsymbol{y} = \{y_1, \cdots, y_n\}$ 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

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

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

Quelques tests

La statistique de test est (toujours) z

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

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

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

```
> qnorm(c(0,.025,.975,1))
[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, σ 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 $\boldsymbol{x} = \{x_1, \cdots, x_n\}$ de loi $\mathcal{N}(\mu, \sigma^2)$.
Pour tester $H_0 : \mu = \mu_0$ contre $H_1 : \mu = \mu_1$, on utilise

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

- 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_ν est la fonction de répartition de la loi de Student $\mathcal{S}\mathrm{td}(\nu)$.

Quelques tests

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

La statistique de test est t

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

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

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

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

```
> qt(c(0,.05), df = 30-1)
[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 (asymmétrique) pour μ, l'espérance de Y

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

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

```
> t.test(x, mu=0, alternative = "less")

  One Sample t-test

data:  y
t = 0.48873, df = 29, p-value = 0.6856
alternative hypothesis: true mean is less than 0
95 percent confidence interval:
      -Inf 0.3691359
```

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

Quelques tests

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

```
> t.test(x, mu=0, alternative = "greater")

  One Sample t-test

data:  x
t = 0.48873, df = 29, p-value = 0.3144
alternative hypothesis: true mean is greater than 0
95 percent confidence interval:
 -0.2042196        Inf
sample estimates:
 mean of x
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 $\boldsymbol{x} = \{x_1, \cdots, x_n\}$ de loi $\mathcal{N}(\mu, \sigma^2)$.
Pour tester $H_0 : \mu = \mu_0$ contre $H_1 : \mu \neq \mu_0$, on utilise

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

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

où T_ν est la fonction de répartition de la loi de Student $\mathcal{S}\text{td}(\nu)$.

Quelques tests

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

```
> t.test(x, mu=0, alternative = "two.sided")

  One Sample t-test

data:  y
t = 0.48873, df = 29, p-value = 0.6287
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.2626142  0.4275306
```

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

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

```
> qt(.975,df=29)
[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.

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

Quelques tests

```
> t.test(blood_pressure$mmhg, mu=140, alternative="
    less")

  One Sample t-test

data:  blood_pressure$mmhg
t = -3.8693, df = 54, p-value = 0.0001481
alternative hypothesis: true mean is less than 140
95 percent confidence interval:
     -Inf 134.3253
sample estimates:
mean of x
      130
```

Quelques tests

Modèle Gaussien avec variances inconnues (mais égales)

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

Soient $\boldsymbol{x} = \{x_1, \cdots, x_m\}$ de loi $\mathcal{N}(\mu_x, \sigma_x^2)$ et $\boldsymbol{y} = \{y_1, \cdots, y_n\}$ 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

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

- 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 utiiser la fonction `t.test`.

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

```
> t.test(x, y, mu=0, alternative = "less")

  Welch Two Sample t-test

data:  x and y
t = -0.37905, df = 37.551, p-value = 0.3534
alternative hypothesis: true difference in means is
    less than 0
95 percent confidence interval:
      -Inf 0.3749006
sample estimates:
 mean of x  mean of y
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 utiiser la fonction `t.test`.

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

```
> t.test(x, y, mu=0, alternative = "greater")

  Welch Two Sample t-test

data:  x and y
t = -0.37905, df = 37.551, p-value = 0.6466
alternative hypothesis: true difference in means is
    greater than 0
95 percent confidence interval:
 -0.5922846        Inf
sample estimates:
 mean of x  mean of y
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 σ_x^2 et σ_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 $H_1 : \mu_x - \mu_y \neq \mu_0$, $\mathcal{N}(\mu_\cdot, \sigma^2)$

Soient $\boldsymbol{x} = \{x_1, \cdots, x_m\}$ de loi $\mathcal{N}(\mu_x, \sigma^2)$ et $\boldsymbol{y} = \{y_1, \cdots, y_n\}$ de loi $\mathcal{N}(\mu_y, \sigma^2)$ (avec la même variance σ^2).
Pour tester $H_0 : \mu_x - \mu_y = \mu_0$ contre $H_1 : \mu_x - \mu_y \neq \mu_0$, on utilise

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

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

Quelques tests

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

```
> set.seed(1)
> x = rnorm(30,0,1)
> y = rnorm(20,0,1)
> t.test(x, y, mu=0, alternative = "two.sided", var.
    equal = TRUE)

  Two Sample t-test

data:  x and y
t = -0.18554, df = 48, p-value = 0.8536
alternative hypothesis: true difference in means is
    not equal to 0
95 percent confidence interval:
 -0.5323591  0.4424086
sample estimates:
 mean of x  mean of y
0.08245817 0.12743344
```

Quelques tests

Sur les données de pression artérielle

```
> t.test(blood_pressure$mmhg ~ blood_pressure$status,
    mu=0, alternative="two.sided", var.equal=TRUE)

  Two Sample t-test

data:  blood_pressure$mmhg by blood_pressure$status
t = -10.468, df = 53, p-value = 1.66e-14
alternative hypothesis: true difference in means
    between group 0 and group 1 is not equal to 0
95 percent confidence interval:
 -37.31328 -25.31339
sample estimates:
mean in group 0 mean in group 1
       112.9200        144.2333
```

Quelques tests

```
> t.test(blood_pressure$mmhg ~ blood_pressure$status)

  Welch Two Sample t-test

data:  blood_pressure$mmhg by blood_pressure$status
t = -10.451, df = 50.886, p-value = 2.887e-14
alternative hypothesis: true difference in means
   between group 0 and group 1 is not equal to 0
95 percent confidence interval:
 -37.32904 -25.29763
sample estimates:
mean in group 0 mean in group 1
       112.9200        144.2333
```

Quelques tests

```
> t.test(blood_pressure$mmhg ~ blood_pressure$status,
    mu=0, alternative="two.sided", var.equal=FALSE)

  Welch Two Sample t-test

data:  blood_pressure$mmhg by blood_pressure$status
t = -10.451, df = 50.886, p-value = 2.887e-14
alternative hypothesis: true difference in means
    between group 0 and group 1 is not equal to 0
95 percent confidence interval:
 -37.32904 -25.29763
sample estimates:
mean in group 0 mean in group 1
       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 $\boldsymbol{x} = \{x_1, \cdots, x_n\}$ 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

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

- 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_ν 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 $H_1 : \sigma^2 > 1$

```
> EnvStats::varTest(x, alternative="greater", conf.
    level = 0.95, sigma.squared = 1)
$statistic
Chi-Squared
   24.76598

$p.value
[1] 0.6903324

$estimate
 variance
0.8539993

$conf.int
      LCL         UCL
0.5819489         Inf
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 $H_1 : \sigma^2 < 1$

```
> EnvStats::varTest(x, alternative="less", conf.level = 0.95, sigma.squared = 1)
$statistic
Chi-Squared
   24.76598

$p.value
[1] 0.3096676

$estimate
 variance
0.8539993

$conf.int
     LCL      UCL
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 $\boldsymbol{x} = \{x_1, \cdots, x_n\}$ 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

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

- 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_ν 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 $H_1 : \sigma^2 \neq 1$

```
> EnvStats::varTest(x, alternative="two.sided", conf.
    level = 0.95, sigma.squared = 1)
$statistic
Chi-Squared
   24.76598

$p.value
[1] 0.6193352

$estimate
 variance
0.8539993

$conf.int
     LCL      UCL
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 $\boldsymbol{x} = \{x_1, \cdots, x_m\}$ de loi $\mathcal{N}(\mu_x, \sigma_x^2)$ et $\boldsymbol{y} = \{y_1, \cdots, y_n\}$ 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

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

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

où F_{ν_1,ν_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`,

```
> set.seed(1)
> x = rnorm(30,0,1)
> y = rnorm(20,0,1)
> var.test(x, y, ratio = 1, alternative = "two.sided")

  F test to compare two variances

data:  x and y
F = 1.7871, num df = 29, denom df = 19, p-value = 0.1896
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.7440377 3.9875896
sample estimates:
ratio of variances
          1.787136
```

Quelques tests

Sur les données de pression artérielle

```
> var.test(blood_pressure$mmhg ~ blood_pressure$status
    ,alternative="two.sided")

  F test to compare two variances

data:  blood_pressure$mmhg by blood_pressure$status
F = 1.0363, num df = 24, denom df = 29, p-value =
    0.918
alternative hypothesis: true ratio of variances is not
     equal to 1
95 percent confidence interval:
 0.4811235 2.2980313
sample estimates:
ratio of variances
          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 appareillées, il est possible que les observations ne soient pas indépendantes entre les deux groupes.

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

Notons $r = \dfrac{\displaystyle\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\displaystyle\sum_{i=1}^{n}(x_i - \overline{x})^2 \cdot \sum_{i=1}^{n}(y_i - \overline{y})^2}}$ 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 $\{(x_1, y_1), \cdots, (x_n, y_n)\}$ 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

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

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, $\text{Var}(x) \sim \text{Var}(y) + 10$

```
> t.test(iq$IQ1,iq$IQ2,mu=-10,alternative="two.sided",
    paired=TRUE)

  Paired t-test

data:  iq$IQ1 and iq$IQ2
t = -1.2854, df = 19, p-value = 0.2141
alternative hypothesis: true difference in means is
    not equal to -10
95 percent confidence interval:
 -11.051338  -9.748662
sample estimates:
mean of the differences
                  -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 $\{(x_1, y_1), \cdots, (x_n, y_n)\}$ 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

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

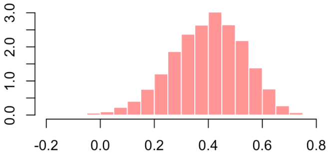
Si $H_0 : \sigma_x^2 = \sigma_y^2$ est vraie, T suit une loi $\mathcal{S}td(n-2)$.

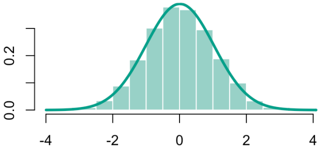
- on rejette H_0 si $|t| > T_{n-2}^{-1}(1-\alpha/2)$

Quelques tests

On peut faire des simulations pour vérifier

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





Quelques tests

Ici,

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

Test d'ajustement de loi

PP Plot, $X \sim F$

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

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

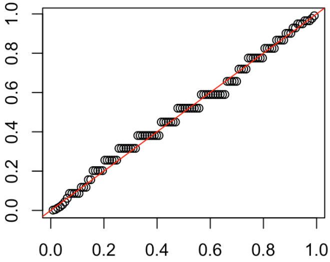
Test $H_0 : X \sim F$ contre $H_1 : X \nsim F$, PP Plot

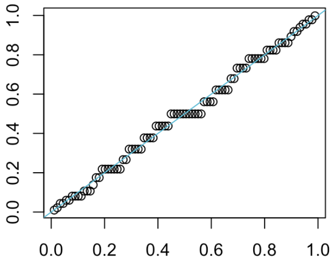
Soit $\{x_1, \cdots, x_n\}$ un échantillon, notons $\{x_{(1)}, \cdots, x_{(n)}\}$ la version ordonnée. Si $H_0 : X \sim F$ est vraie, les points sont alignés sur la première bissectrice $(y = x)$.

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 $\mathcal{N}(\overline{x}, \widehat{s}^2)$

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





Test d'ajustement de loi

QQ Plot, $X \sim F$

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

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

Test $H_0 : X \sim F$ contre $H_1 : X \nsim F$, QQ Plot

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

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

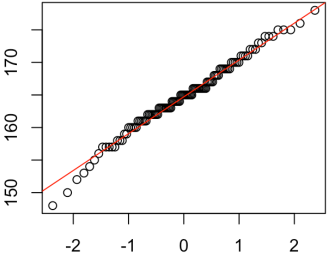
Test $H_0 : X \sim \mathcal{N}$ contre $H_1 : X \nsim \mathcal{N}$, QQ Plot Gaussien

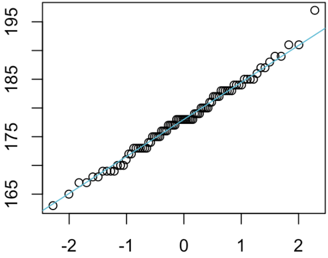
Soit $\{x_1, \cdots, x_n\}$ un échantillon, notons $\{x_{(1)}, \cdots, x_{(n)}\}$ la version ordonnée. Si $H_0 : X \sim \mathcal{N}(\mu, \sigma^2)$ est vraie, les points sont alignés sur la droite $y = \mu + \sigma x$.

Test d'ajustement de loi

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

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

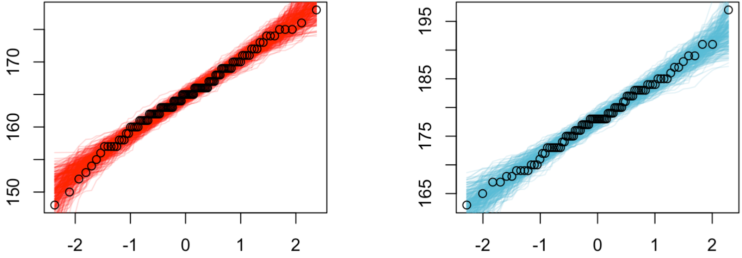




Test d'ajustement de loi

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

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



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 $\boldsymbol{x} = \{x_1, \cdots, x_n\}$ de loi F_X.
Pour tester $H_0 : F_X = F$ contre $H_1 : F_X \neq F$, on utilise

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

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

où K_ν est la loi du test de Komogorov-Smirnov.

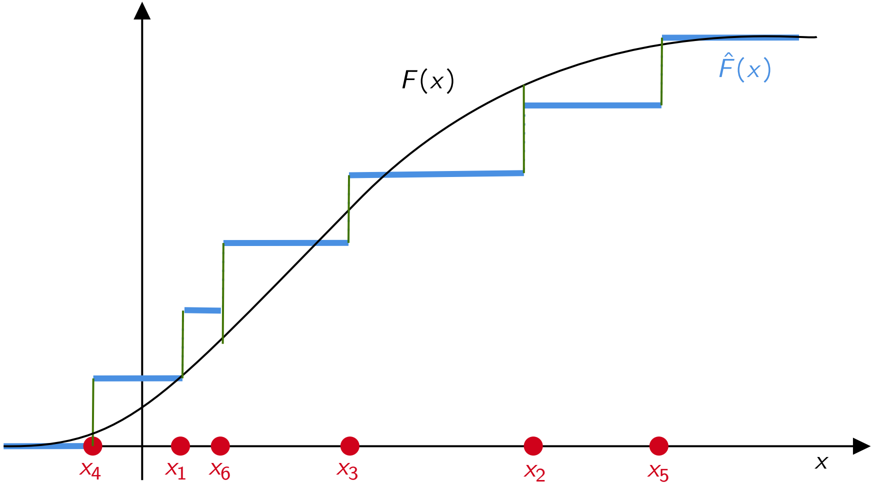
Statistique Kolmogorov-Smirnov

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

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

Test d'ajustement de loi

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



Test d'ajustement de loi

Numériquement, on utilise `ks.test` sous R

```
> ks.test(x, "pnorm", mean = 165, sd =5)

  One-sample Kolmogorov-Smirnov test

data:  x
D = 0.081455, p-value = 0.4472
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 $H_0 : X \sim \mathcal{N}(\overline{x}, s^2)$, 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 $H_0 : X \sim \mathcal{N}(\overline{x}, s^2)$,

```
> ks.test(x,"pnorm",mean=mean(x),sd=sd(x))

  One-sample Kolmogorov-Smirnov test

data:  x
D = 0.070851, p-value = 0.6275
alternative hypothesis: two-sided
```

```
> library(nortest)
> lillie.test(x)

  Lilliefors (Kolmogorov-Smirnov) normality test

data:  x
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 ★★★

```
> ad.test(x)

  Anderson-Darling normality test

data:  x
A = 0.38576, p-value = 0.3858
> sf.test(x)

  Shapiro-Francia normality test

data:  x
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 $\boldsymbol{x} = \{x_1, \cdots, x_m\}$ de loi F_X et $\boldsymbol{y} = \{y_1, \cdots, y_n\}$ de loi F_Y.

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

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

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

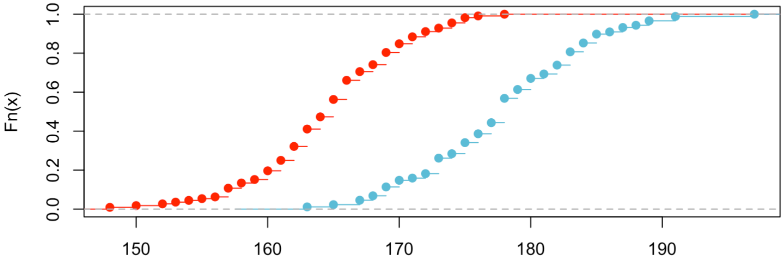
où K_{ν_1,ν_2} est la loi du test de Komogorov-Smirnov.

Test d'ajustement de loi

```
> ks.test(x,y)

  Two-sample Kolmogorov-Smirnov test

data:  x and y
D = 0.7289, p-value < 2.2e-16
alternative hypothesis: two-sided
> plot(ecdf(x), xlim = range(c(x,y)))
> 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 $\boldsymbol{x} = \{x_1, \cdots, x_m\}$ de loi F_X et $\boldsymbol{y} = \{y_1, \cdots, y_n\}$ 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

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

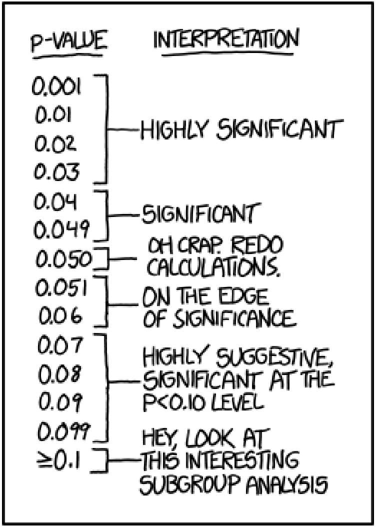
- on rejette H_0 si $d > c(\alpha)\sqrt{\dfrac{n+m}{n\cdot m}}$

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



Exemple de tests

EXEMPLE 1

Un chercheur émet l'hypothèse que l'âge moyen des femmes à leur premier mariage a augmenté depuis la dernière étude menée sur le sujet en 2011, qui avait établi la moyenne à 31 ans. Formuler les hypothèses H_0 et H_1.

Source: Institut de la statistique du Québec. *Les mariages au Québec en 2011*, juin 2014.

(via Simard (2015))

On pourrait écrire $H_0 : \mu = \mu_0 = 31$ et $H_1 : \mu > \mu_0 = 31$.

EXEMPLE 2

Une association de consommateurs examine un échantillon de 100 contenants de sirop d'érable pour vérifier si le volume moyen de sirop dans les contenants est bien de 540 ml, comme l'indique l'étiquette. Formuler les hypothèses H_0 et H_1.

(via Simard (2015))

On pourrait écrire $H_0 : \mu = \mu_0 = 540$ et $H_1 : \mu < \mu_0 = 540$.
(L'association pourrait vouloir démontrer que le consommateur est victime d'une publicité mensongère)

Exemple de tests

EXEMPLE 3

Un producteur de sirop d'érable prélève un échantillon de 100 contenants de sirop dans la production d'une journée afin de s'assurer que le volume moyen de sirop est bien égal à 540 ml. Formuler les hypothèses H_0 et H_1.

(via Simard (2015))

On pourrait écrire $H_0 : \mu = \mu_0 = 540$ et $H_1 : \mu \neq \mu_0 = 540$.
(Le producteur veut s'assurer que la moyenne n'est ni inférieure ni supérieure à 540 ml)

Exemple de tests

EXEMPLE 2

On veut tester la durée, en kilomètres, d'une nouvelle semelle de pneus de voiture. Une analyse échantillonnale de 12 pneus a donné une durée moyenne de 53 870 km avec un écart type corrigé de 7 760 km. Au seuil de signification de 0,01, peut-on dire que la nouvelle semelle améliore la durée moyenne des pneus actuels, si celle-ci suit une distribution normale dont la moyenne est de 50 000 km ?

(via Simard (2015))

On veut tester ici $H_0 : \mu = \mu_0$ contre $H_1 : \mu > \mu_0$ avec $\mu_0 = 50000$ km. Donc la région de rejet (de H_0) sera de la forme "$\overline{x}$ grand" (ou $\overline{x} > k$).

On va supposer que la durée d'un pneu suit une loi normale $\mathcal{N}(\mu, \sigma^2)$. On a un échantillon de taille $n = 12$, et on nous dit que $\overline{x} = 53870$ et $s = 7760$.

Si H_0 est vraie,

$$T = \sqrt{12}\frac{\overline{X} - \mu_0}{s} \sim \mathcal{S}td(12 - 1).$$

Exemple de tests

EXEMPLE 2

On veut tester la durée, en kilomètres, d'une nouvelle semelle de pneus de voiture. Une analyse échantillonnale de 12 pneus a donné une durée moyenne de 53 870 km avec un écart type corrigé de 7 760 km. Au seuil de signification de 0,01, peut-on dire que la nouvelle semelle améliore la durée moyenne des pneus actuels, si celle-ci suit une distribution normale dont la moyenne est de 50 000 km ?

(via Simard (2015))

Si H_0 est vraie,

$$\mathbb{P}\left(T = \sqrt{12}\frac{\overline{X} - \mu_0}{s} > \sqrt{12}\frac{k - \mu_0}{s} \middle| T \sim \mathcal{S}td(12-1)\right) = \alpha = 1\%$$

donc

$$\sqrt{12}\frac{k - \mu_0}{s} = t_{12-1,99\%} \text{ ou } k = \mu_0 + \frac{s}{\sqrt{12}} t_{12-1,99\%}$$

```
> 50000 + qt(.99,12-1)*7760/sqrt(12)
[1] 56088.82
```

Exemple de tests

EXEMPLE 2

On veut tester la durée, en kilomètres, d'une nouvelle semelle de pneus de voiture. Une analyse échantillonnale de 12 pneus a donné une durée moyenne de 53 870 km avec un écart type corrigé de 7 760 km. Au seuil de signification de 0,01, peut-on dire que la nouvelle semelle améliore la durée moyenne des pneus actuels, si celle-ci suit une distribution normale dont la moyenne est de 50 000 km ?

(via Simard (2015))

- si $\overline{x} > 56088.82$ on rejette $H_0 : \mu = 50000$ (au profit de $H_1 : \mu > 50000$)
- si $\overline{x} < 56088.82$ on rejette $H_1 : \mu > 50000$ (au profit de $H_0 : \mu = 50000$)

Comme $\overline{x} = 53870 < 56088$, on ne rejette pas H_0.

Peut-on dire que la nouvelle semelle améliore la durée moyenne des pneus actuels : non, on ne peut pas le dire.

Exemple de tests

EXEMPLE

Seulement 20 % des clients d'un magasin acquittent leurs achats par paiement direct. Le propriétaire du magasin organise une campagne de promotion afin d'inciter un plus grand nombre de clients à employer ce mode de paiement. Quelque temps après la fin de la campagne, on veut en vérifier l'efficacité. Dans un échantillon de 150 clients, 42 ont utilisé le paiement direct. Peut-on accepter l'hypothèse selon laquelle la campagne de promotion a été efficace ? Faire un test au seuil de signification de 0,01.

(via Simard (2015))

On a ici un test de proportion de la forme $H_0 : p = p_0 = 20\%$ contre $H_1 : p > p_0 = 20\%$.

On dispose de $n = 150$ observations et $\widehat{p} = \overline{y} = 42/150 = 28\%$

On peut utiliser un test Gaussien (avec une approximation normale) en notant que pour tester $H_0 : p = p_0 = 20\%$ contre $H_1 : p > p_0 = 20\%$, on rejette H_0 si $\overline{x} > k$). Si H_0 est vraie,

$$Z = \sqrt{150}\frac{\overline{X} - p_0}{\sqrt{p_0(1-p_0)}} \approx \mathcal{N}(0,1).$$

Exemple de tests

EXEMPLE

Seulement 20 % des clients d'un magasin acquittent leurs achats par paiement direct. Le propriétaire du magasin organise une campagne de promotion afin d'inciter un plus grand nombre de clients à employer ce mode de paiement. Quelque temps après la fin de la campagne, on veut en vérifier l'efficacité. Dans un échantillon de 150 clients, 42 ont utilisé le paiement direct. Peut-on accepter l'hypothèse selon laquelle la campagne de promotion a été efficace ? Faire un test au seuil de signification de 0,01.

(via Simard (2015))

Si H_0 est vraie,

$$\mathbb{P}\left(Z=\sqrt{150}\frac{\overline{X}-p_0}{\sqrt{p_0(1-p_0)}}>\sqrt{150}\frac{k-p_0}{\sqrt{p_0(1-p_0)}}\middle|Z\sim\mathcal{N}(0,1)\right)=\alpha=1\%$$

donc

$$\sqrt{150}\frac{k-p_0}{\sqrt{p_0(1-p_0)}}=u_{99\%}\text{ ou }k=p_0+\frac{\sqrt{p_0(1-p_0)}}{\sqrt{150}}u_{99\%}$$

```
> 0.2 + qnorm(.99)*sqrt(0.2*0.8)/sqrt(150)
[1] 0.2759782
```

Exemple de tests

EXEMPLE

Seulement 20 % des clients d'un magasin acquittent leurs achats par paiement direct. Le propriétaire du magasin organise une campagne de promotion afin d'inciter un plus grand nombre de clients à employer ce mode de paiement. Quelque temps après la fin de la campagne, on veut en vérifier l'efficacité. Dans un échantillon de 150 clients, 42 ont utilisé le paiement direct. Peut-on accepter l'hypothèse selon laquelle la campagne de promotion a été efficace ? Faire un test au seuil de signification de 0,01.

(via Simard (2015))

- si $\overline{x} > 27.6\%$ on rejette $H_0 : p = 20\%$ (au profit de $H_1 : p > 20\%$)
- si $\overline{x} < 27.6\%$ on rejette $H_1 : p > 20\%$ (au profit de $H_0 : p = 20\%$)

Comme $\overline{x} = 28\% > 27.6\%$, on rejette H_0.

Peut-on accepter l'hypothèse selon laquelle la campagne de promotion a été efficace ? : oui, on l'accepte

Exemple de tests

EXEMPLE

Seulement 20 % des clients d'un magasin acquittent leurs achats par paiement direct. Le propriétaire du magasin organise une campagne de promotion afin d'inciter un plus grand nombre de clients à employer ce mode de paiement. Quelque temps après la fin de la campagne, on veut en vérifier l'efficacité. Dans un échantillon de 150 clients, 42 ont utilisé le paiement direct. Peut-on accepter l'hypothèse selon laquelle la campagne de promotion a été efficace ? Faire un test au seuil de signification de 0,01.

(via Simard (2015))

On peut faire le test directement

```
> prop.test(42,150,.2,alternative = "greater")

  1-sample proportions test with continuity correction

data:  42 out of 150, null probability 0.2
X-squared = 5.5104, df = 1, p-value = 0.009452
alternative hypothesis: true p is greater than 0.2
95 percent confidence interval:
 0.2209469 1.0000000
```

La p-value vaut $0.009452 < 1\%$ donc on rejette H_0.

Exemple de tests

EXEMPLE

Dans certaines succursales d'une chaîne de restauration rapide, on utilise un afficheur électronique pour présenter des photos de divers produits. Pour mesurer l'efficacité de ce type de promotion, on prélève un échantillon de 36 jours, et on compare la moyenne quotidienne des ventes du produit vedette de deux succursales, l'une disposant d'un afficheur électronique (A) et l'autre pas (A′). Pour ces 36 jours, la moyenne quotidienne des ventes a été de 170 unités avec un écart type corrigé de 6 unités pour le restaurant disposant d'un afficheur et de 165 unités avec un écart type corrigé de 5 unités pour l'autre restaurant. Au seuil de signification de 0,05, peut-on affirmer que la moyenne quotidienne des ventes du produit vedette est plus élevée pour la succursale utilisant un afficheur électronique ?

(via Simard (2015))

On a ici un test de comparaison de deux moyennes, de la forme $H_0 : \mu_A - \mu_{A'} = 0$ contre $H_1 : \mu_A - \mu_{A'} > 0$, avec des échantillons de même taille ($m = n = 36$ a priori non appareillés).

Avec les notations usuelles, $\overline{x} = 170$ avec $s_x = 6$, et $\overline{y} = 165$ avec $s_y = 5$. Si H_0 est vraie, alors

$$\Delta = \overline{X} - \overline{Y} \sim \mathcal{N}\left(0, \frac{s_x^2}{36} + \frac{s_y^2}{36}\right) \text{ i.e. } \mathcal{N}(0, 1.3^2)$$

Exemple de tests

EXEMPLE

> Dans certaines succursales d'une chaîne de restauration rapide, on utilise un afficheur électronique pour présenter des photos de divers produits. Pour mesurer l'efficacité de ce type de promotion, on prélève un échantillon de 36 jours, et on compare la moyenne quotidienne des ventes du produit vedette de deux succursales, l'une disposant d'un afficheur électronique (A) et l'autre pas (A′). Pour ces 36 jours, la moyenne quotidienne des ventes a été de 170 unités avec un écart type corrigé de 6 unités pour le restaurant disposant d'un afficheur et de 165 unités avec un écart type corrigé de 5 unités pour l'autre restaurant. Au seuil de signification de 0,05, peut-on affirmer que la moyenne quotidienne des ventes du produit vedette est plus élevée pour la succursale utilisant un afficheur électronique ?

(via Simard (2015))

Comme on veut tester $H_0 : \mu_A - \mu_{A'} = 0$ contre $H_1 : \mu_A - \mu_{A'} > 0$, on va rejeter H_0 si $\delta = \overline{x} - \overline{y}$ est "grand", i.e. $\delta = k$.

$$\mathbb{P}\big[\Delta > k \big| \Delta \sim \mathcal{N}(0, 1.3^2)\big] = 95\%$$

soit $k = u_{95\%} \times 1.3 = 1.645 \times 1.3 = 2.138$

```
> qnorm(.95,0,1.3)
[1] 2.13831
> qnorm(.95)*1.3
[1] 2.13831
```

Exemple de tests

EXEMPLE

Dans certaines succursales d'une chaîne de restauration rapide, on utilise un afficheur électronique pour présenter des photos de divers produits. Pour mesurer l'efficacité de ce type de promotion, on prélève un échantillon de 36 jours, et on compare la moyenne quotidienne des ventes du produit vedette de deux succursales, l'une disposant d'un afficheur électronique (A) et l'autre pas (A′). Pour ces 36 jours, la moyenne quotidienne des ventes a été de 170 unités avec un écart type corrigé de 6 unités pour le restaurant disposant d'un afficheur et de 165 unités avec un écart type corrigé de 5 unités pour l'autre restaurant. Au seuil de signification de 0,05, peut-on affirmer que la moyenne quotidienne des ventes du produit vedette est plus élevée pour la succursale utilisant un afficheur électronique ?

(via Simard (2015))

- si $\overline{x}-\overline{y}>2.13$ on rejette $H_0:\mu_A-\mu_{A'}=0$ (au profit de $H_1:\mu_A-\mu_{A'}>0$)
- si $\overline{x}-\overline{y}<2.13$ on rejette $H_1:\mu_A-\mu_{A'}>0$ (au profit de $H_0:\mu_A-\mu_{A'}=0$)

Comme $\overline{x}-\overline{y}=5>2.13$, on rejette H_0.

Au seuil de signification de 0.05, peut-on affirmer que la moyenne quotidienne des ventes du produit vedette est plus élevée pour la succursale utilisant un afficheur électronique ? : oui.

Exemple de tests

EXEMPLE

Dans le cadre d'une étude sur le fonctionnement du système nerveux, on a voulu savoir si l'écoute de la musique modifie le temps de réaction à un stimulus visuel chez les adolescents. Pour ce faire, on a choisi 8 adolescents au hasard et l'on a mesuré leur temps de réaction à l'apparition d'une image sur un écran. Deux séries de mesures ont été effectuées : une série dans le calme et une autre alors que le sujet écoutait de la musique douce avec un casque. La forme visualisée était un carré bleu et son nombre d'apparitions était fixé à 20. Le tableau ci-dessous donne le temps de réaction moyen des adolescents, en millisecondes (ms). Au seuil de signification de 0,05, peut-on conclure que le temps de réaction chez les adolescents est influencé par l'écoute de la musique ? (On suppose que la variable aléatoire D de la différence entre y et x suit une loi normale.)

Temps de réaction (en ms)

Sujet	1	2	3	4	5	6	7	8
Sans musique (x)	319	262	293	374	270	265	261	303
Avec musique (y)	299	256	312	357	277	279	253	286
$d = y - x$	-20	-6	19	-17	7	14	-8	-17

(via Simard (2015))

On a ici un test de comparaison de deux moyennes, de la forme $H_0 : \mu_x - \mu_y = 0$ contre $H_1 : \mu_x - \mu_y \neq 0$, avec des échantillons de taille $n = 8$.

Exemple de tests

Si on oublie que les données sont appareillées

```
> x = c(319, 262, 293, 374, 270, 265, 261, 303)
> y = c(299, 256, 312, 357, 277, 279, 253, 286)
> t.test(x, y)

  Welch Two Sample t-test

data:  x and y
t = 0.19247, df = 13.698, p-value = 0.8502
alternative hypothesis: true difference in means is
    not equal to 0
95 percent confidence interval:
 -35.5836  42.5836
sample estimates:
mean of x mean of y
  293.375   289.875
```

La p-value vaut 0.8502 > 5% donc on ne rejette pas $H_0 : \mu_x = \mu_y$.

Exemple de tests

Si on tient compte du fait que les données sont appareillées

```
> t.test(x, y, paired=TRUE)

  Paired t-test

data:  x and y
t = 0.65767, df = 7, p-value = 0.5318
alternative hypothesis: true difference in means is
    not equal to 0
95 percent confidence interval:
 -9.084029 16.084029
sample estimates:
mean of the differences
                    3.5
```

La p-value vaut $0.5318 > 5\%$ donc on ne rejette pas $H_0 : \mu_x = \mu_y$.

Exemple de tests

EXEMPLE

Une étude menée auprès d'un échantillon de 450 hommes et 500 femmes indique que 17 % des hommes et 13 % des femmes dorment moins de 6,5 heures par nuit. Au seuil de signification de 0,05, peut-on en conclure que le pourcentage de personnes qui dorment moins de 6,5 heures par nuit est plus élevé chez les hommes que chez les femmes ?

Source : Statistique Canada. *Enquête sociale générale, 2006.*

(via Simard (2015))

On a ici un test de comparaison de deux proportion, de la forme $H_0 : p_x - p_y = 0$ contre $H_1 : p_x - p_y > 00$.

Exemple de tests

```
> prop.test(x = c(450*.17,500*.13), n = c(450, 500), p
    = NULL, alternative = "greater")

  2-sample test for equality of proportions with
    continuity correction

X-squared = 2.6822, df = 1, p-value = 0.05074
alternative hypothesis: greater
95 percent confidence interval:
 -0.0003254031  1.0000000000
sample estimates:
prop 1 prop 2
  0.17   0.13
```

La p-value vaut 0.05074 > 5% donc on ne rejette pas $H_0 : p_x = p_y$ (mais c'est vraiment limite).