Estimation et tests d'hypothèses

Thierry Dhorne

20 septembre 2016

Rappels

État des lieux

- quatre situations étudiées
  - ➢ test sur une probabilité (pas en détail)
  - ➢ test sur une espérance à variance connue (peu d'intérêt pratique)
  - ➢ test sur une variance (à connaître parfaitement)
  - ➢ test sur une espérance à variance inconnue connue (à connaître parfaitement)

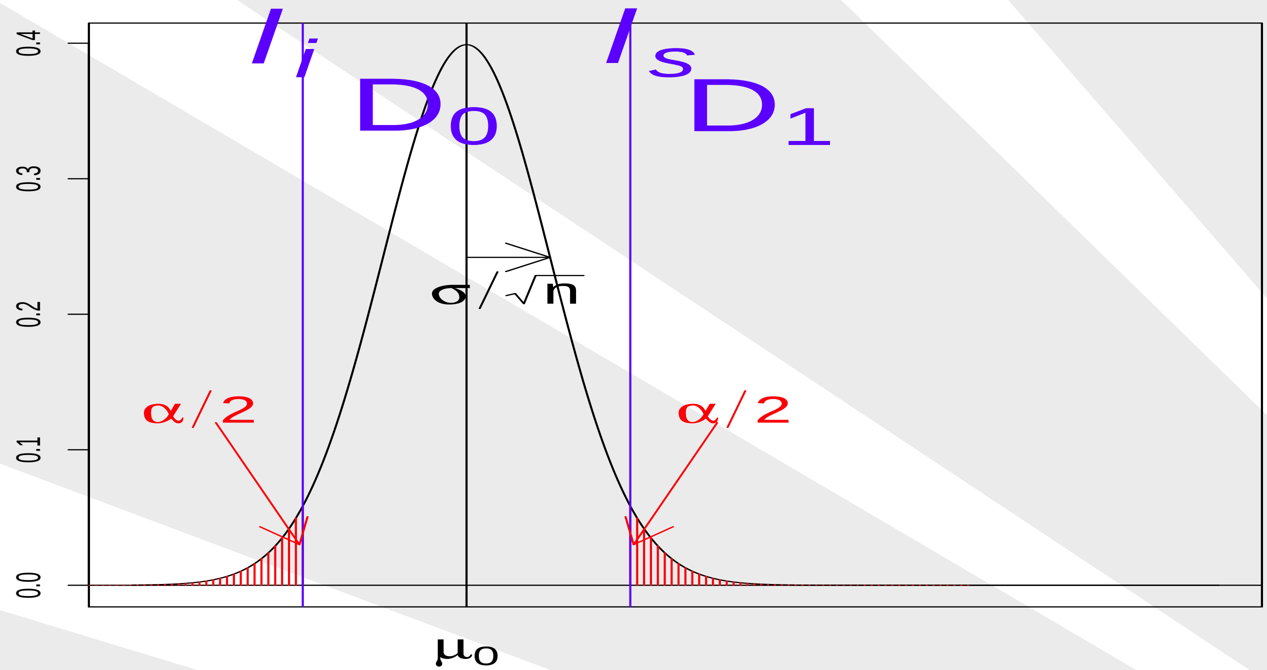
Plan de conception, de mise en œuvre et de validation d'un test (version classique)

0. spécification du modèle : X(à préciser) $\rightsquigarrow \mathscr{L}(\theta_1, \theta_2, ...)$ à définir
1. spécification des hypothèses (en fonction des paramètres du modèle)
   — hypothèse nulle : H_O
   — hypothèse alternative H_1
2. statistique de test (T : v.a. fonction de l'échantillon) et loi de cette statistique (sous H_0 et sous H_1)
3. spécification de la forme des régions de décision et normalisation des régions (risque de 1ère espèce)
4. expérimentation, obtention de l'échantillon, calcul de la statistique de test observée t_o
5. décision
6. éventuellement si $D(H_0)$ calcul de la puissance

Plan de conception, de mise en œuvre et de validation d'un test (version moderne)

0. spécification du modèle : X(à préciser) $\leadsto \mathscr{L}(\theta_1, \theta_2, ...)$ à définir
1. spécification des hypothèses (en fonction des paramètres du modèle)
   - hypothèse nulle : H_O
   - hypothèse alternative H_1
2. statistique de test (T : v.a. fonction de l'échantillon) et loi de cette statistique (sous H_0 et sous H_1)
3. spécification de la forme des régions de décision
4. expérimentation, obtention de l'échantillon, calcul de la statistique de test observée t_o
5. calcul de la probabilité critique $\min_{P(D_1)} P_{t_0 \in D_1}$
6. décision
7. éventuellement si $D(H_0)$ calcul de la puissance

Rappel : normalisation de la région de décision (approche classique)



- la frontière l (en bleu) est telle que

$$P(T > l) = \alpha$$

➤ donc l est fonction du quantile de proba $1 - \frac{\alpha}{2}$ d'une gaussienne selon

$$l = \mu_0 + \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}}$$

Puissance d’un test

Risques

- on connait deux risques :

— le risque de 1ère espèce qui est contrôlé lors de la normalisation (version classique) ou lors de la décision (version moderne)

— le risque de 2ème espèce qui est donc contraint par le choix du risque de 1ère espèce

- le risque de 2ème espèce n'est donc pas contrôlé

➤ il peut être judicieux de le calculer

Intérêt d'une expérience de test

- en général lorsque l'on met en place un test
- ➤ on souhaite décider H_1 lorsque $H1$ est vraie
- ➤ exemple : on veut identifier un médicament s'il est actif
- ➤ exemple : on veut détecter un déréglage si la machine est déréglée
- ★ on souhaite donc que : $P_{H_1}(D(H_1)$ soit élevée
- or $P_{H_1}(D(H_1)) = 1 - P_{H_1}(D(H_0)) = 1 - \beta$ (risque de deuxième espèce)
- minimiser le risque de deuxième espèce est donc équivalent à maximiser $P_{H_1}(D(H_1))$

Puissance d'un test

Puissance d'un test

La puissance d'un test est sa capacité à détecter H_1 (quand H_1 est vraie), elle est donc égale à :

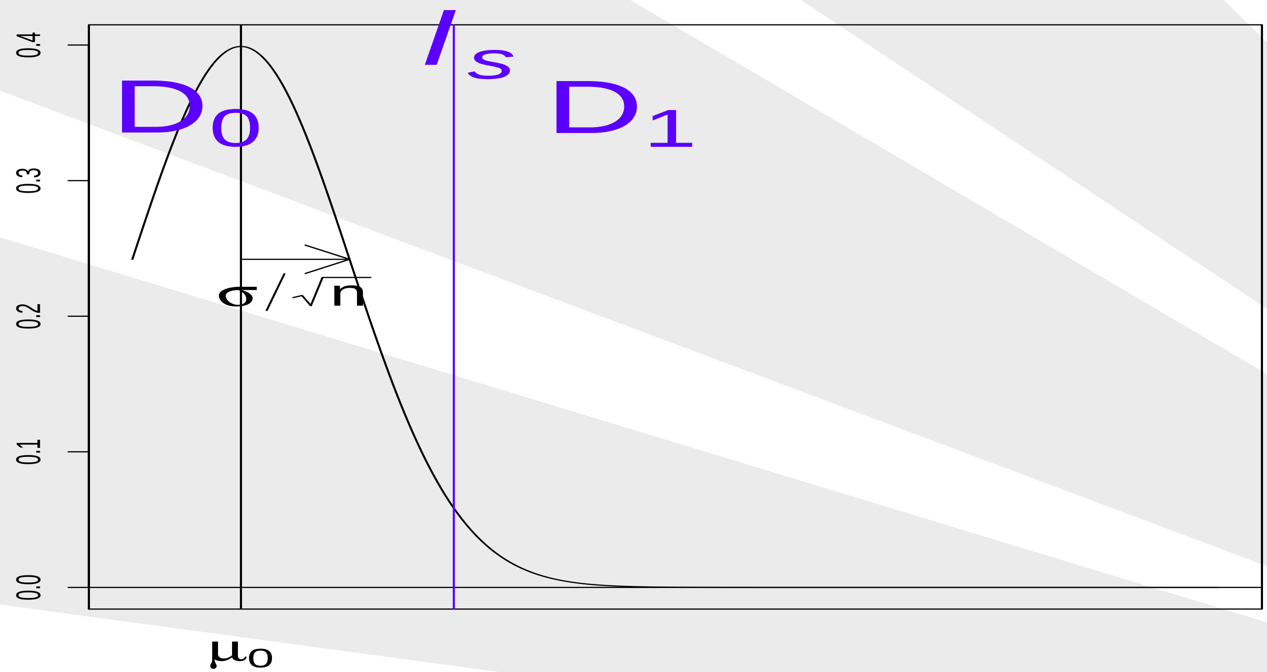
$$P_{H_1}(D(H_1))$$

elle est donc égale à $1 - P_{H_1}(D(H_0))$, c'est-à-dire à $1 - \beta$ où β est le risque de 2ème espèce.

Puissance d'un test : rappel

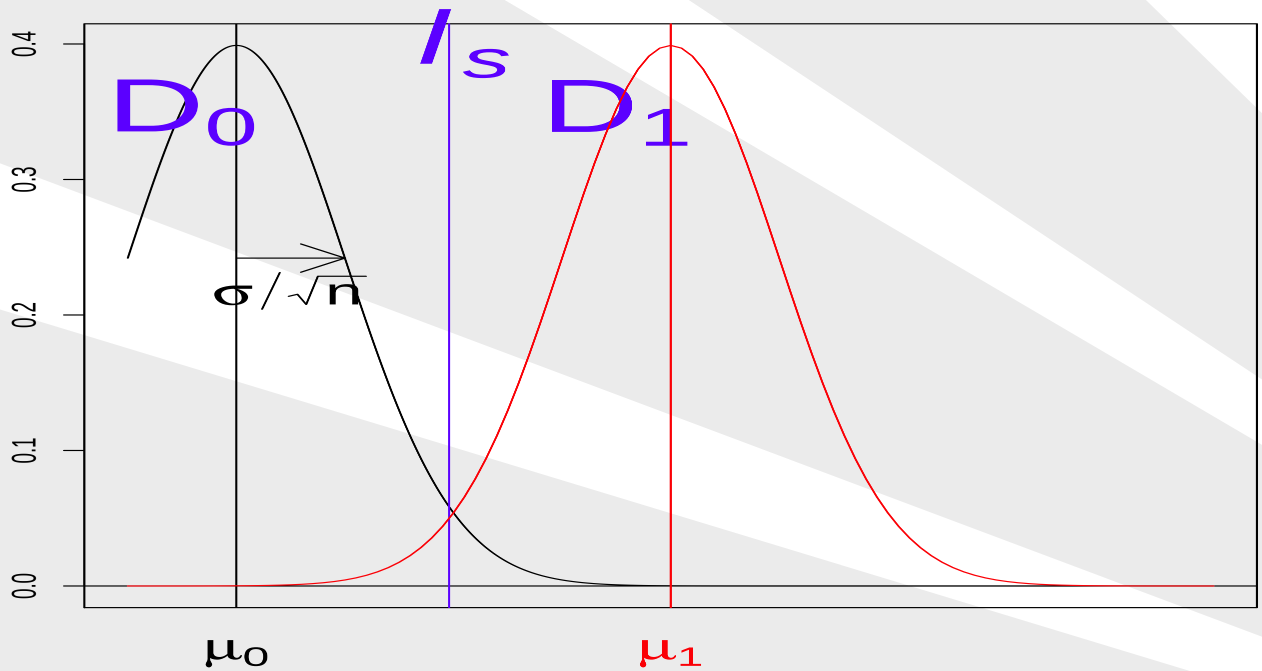
- le calcul de la puissance a été réalisé pour le test sur l'espérance à variance connue

➤ la région de décision est ainsi normalisée (on a « décalé » le graphique)



Puissance d'un test : alternative

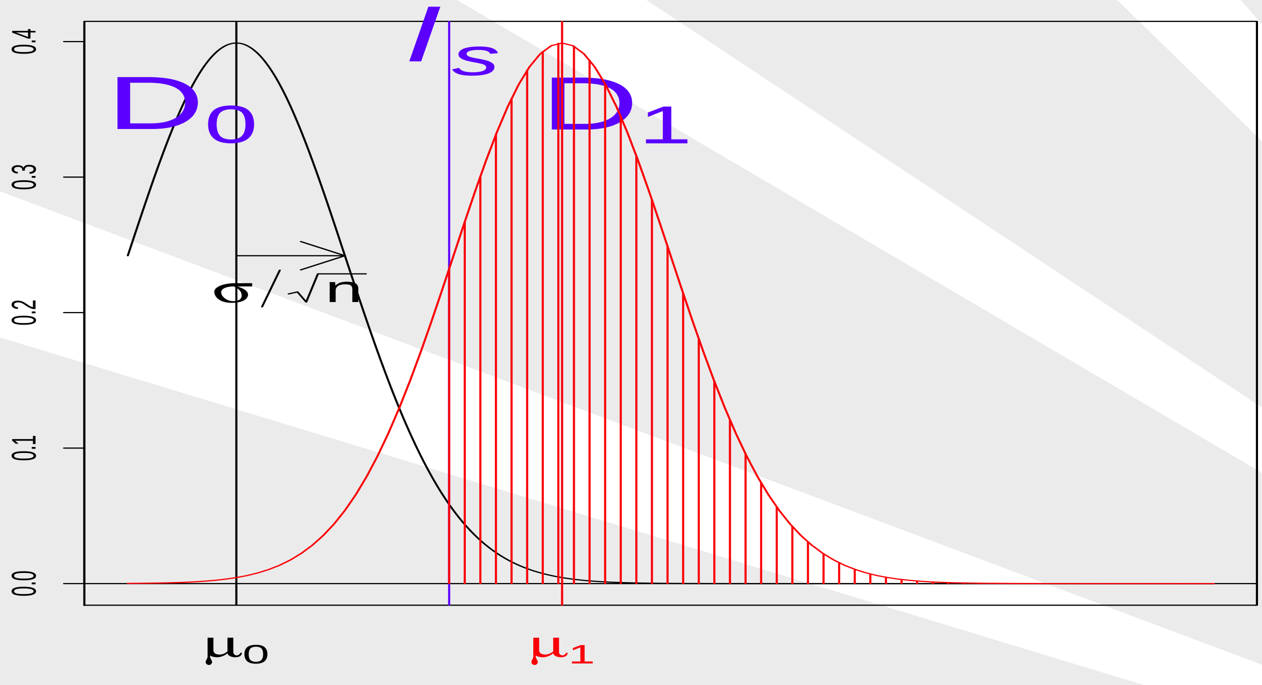
- on représente la loi de la statistique de test pour une valeur quelconque de H_1 par exemple μ_1
  - la courbe associée est en rouge



Puissance d'un test : représentation

- la puissance est la probabilité de décider $H1$ (= être dans D_1) quand H_1 est vraie (= prendre la loi rouge)

➤ cette puissance est donc la surface en rouge



➤ qui vaut donc

$$P(\mathcal{N}(\mu_1, \frac{\sigma}{\sqrt{n}}) > l) = 1 - F_{\mu_1, \frac{\sigma}{\sqrt{n}}}(l)$$

Puissance d'un test : représentation

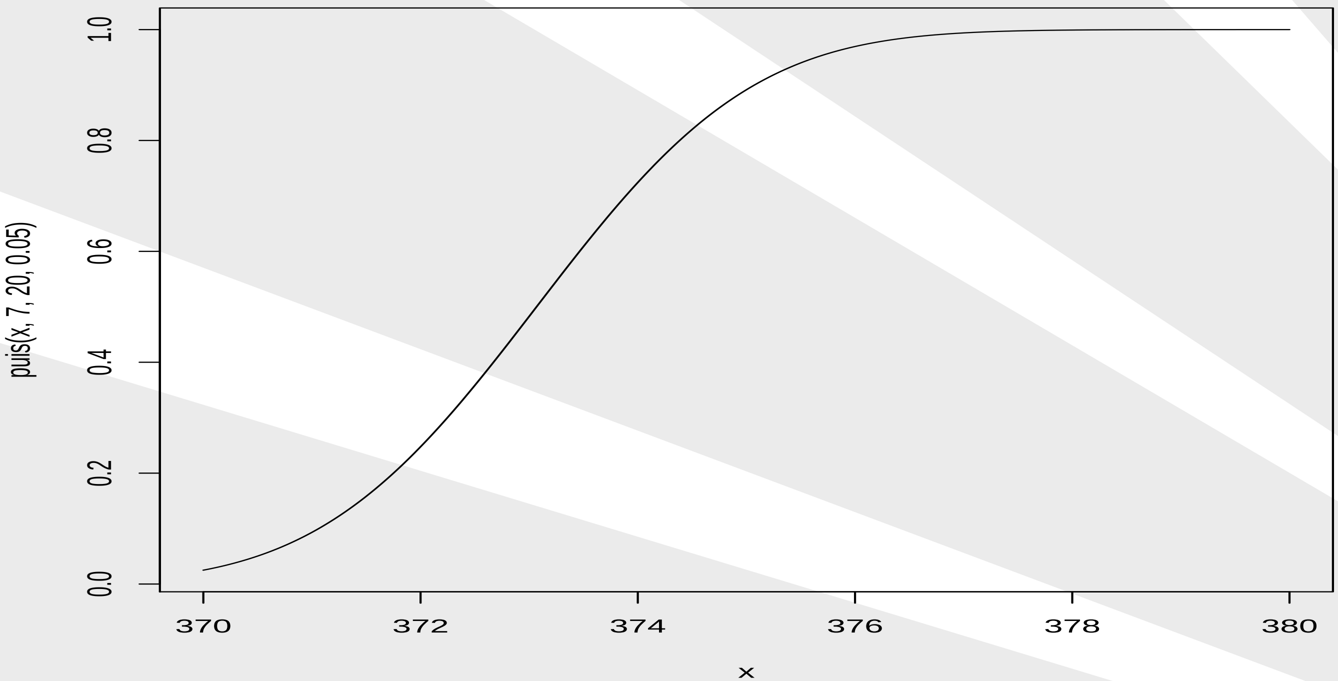
- la puissance vaut donc

$$
\begin{aligned}
1 - F_{\mu_1, \frac{\sigma}{\sqrt{n}}}(l) &= 1 - F_{0,1}\left(\frac{l - \mu_1}{\frac{\sigma}{\sqrt{n}}}\right) \\
&= 1 - F_{0,1}\left(\frac{\mu_0 + \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}} - \mu_1}{\frac{\sigma}{\sqrt{n}}}\right) \\
&= 1 - F_{0,1}\left(\frac{\mu_0 - \mu_1}{\frac{\sigma}{\sqrt{n}}} + z_{1-\frac{\alpha}{2}}\right)
\end{aligned}
$$

Puissance d'un test : contrôle

- la puissance du test dépend

➤ de la valeur de l'alternative (μ_1)

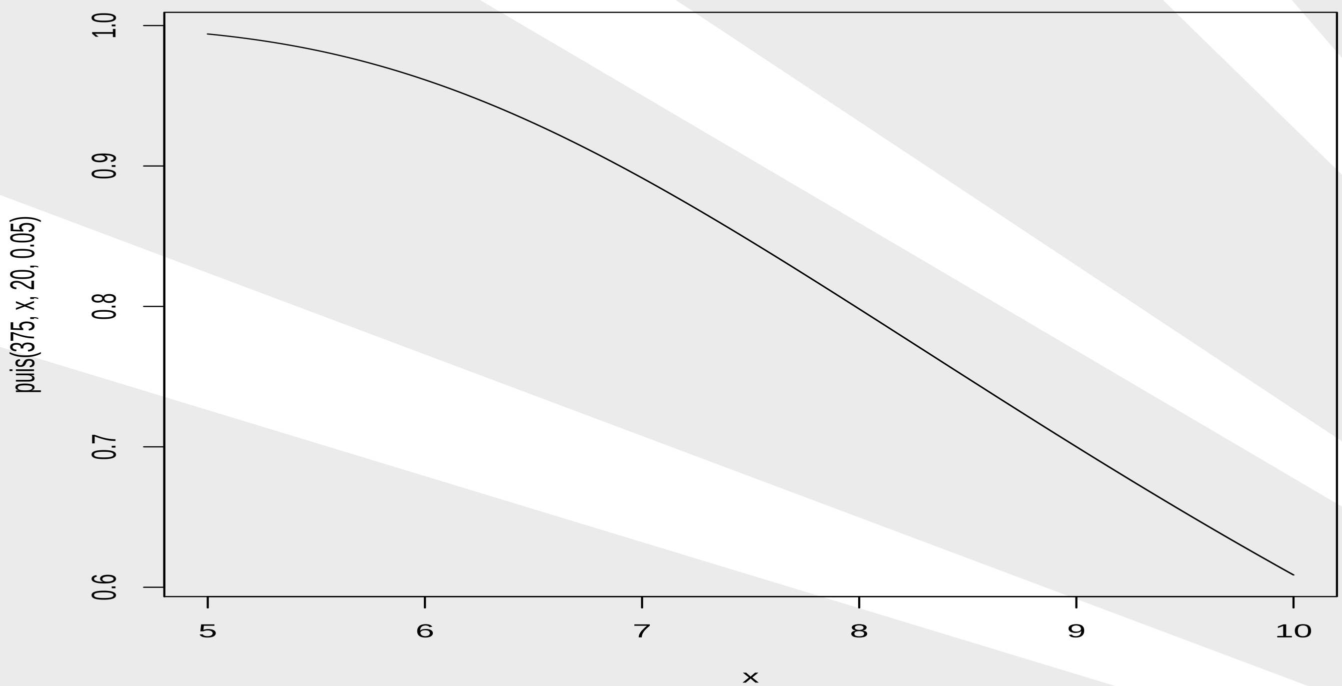


➤ plus $\mu 1$ s'éloigne de μ_0 plus il est facile de détecter la différence

- pour un déréglage de 5 g, la puissance est 0.8915

Puissance d'un test : contrôle

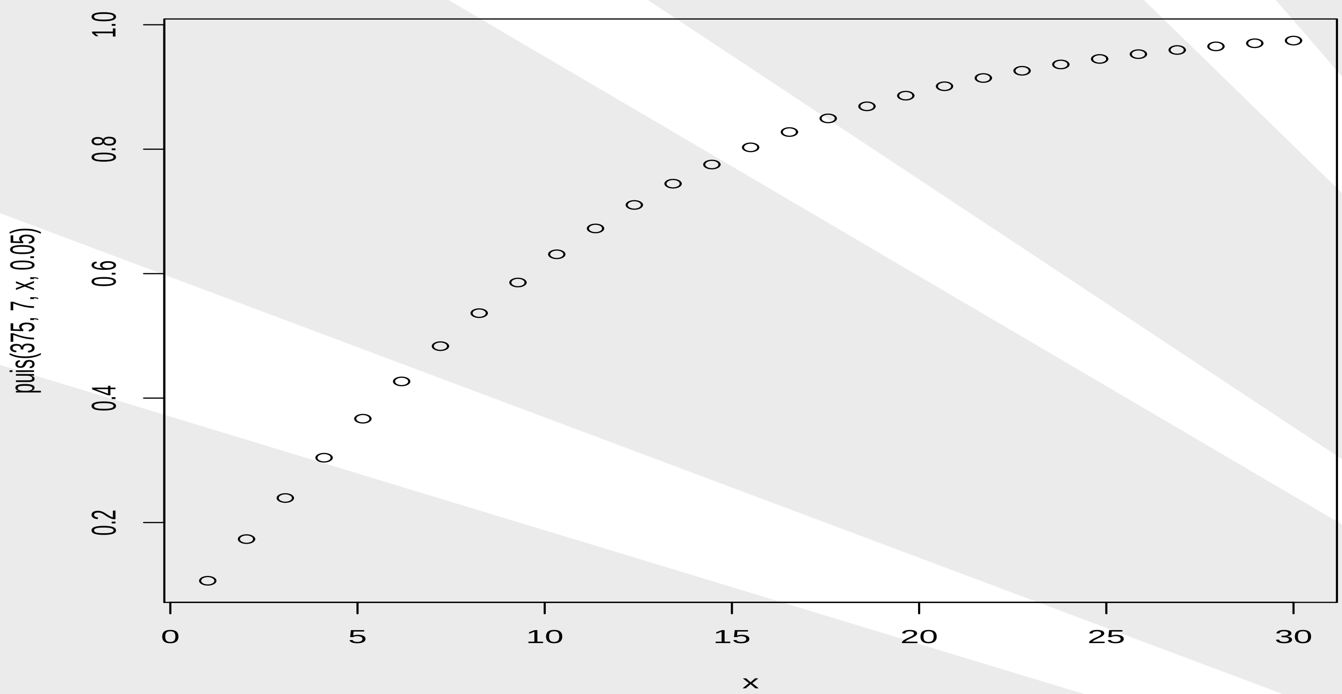
- la puissance du test dépend
  - de la valeur de l'écart-type (σ)



- plus σ augmente plus il est difficile de détecter la différence

Puissance d'un test : contrôle

- la puissance du test dépend
  - ➤ de la taille de l'échantillon n



➤ plus n augmente plus il est facile de détecter la différence

Puissance d'un test : contrôle

- la puissance du test dépend AUSSI
- ➤ du risque de première espèce
- ➤ mais, en toute rigueur, le risque de première espèce doit être fixé a priori
- ★ on ne devrait donc pas « jouer » avec le risque de première espèce

Cas d'une variance inconnue

Puissance d'un test : cas d'une variance inconnue

- le calcul de la puissance a été réalisé pour le test sur l'espérance à variance connue
- on le réalise maintenant lorsque la variance n'est pas connue et qu'on l'estime

★ on est alors obligé d'utiliser une v.a. centrée (sur μ_0) et pseudo-réduite

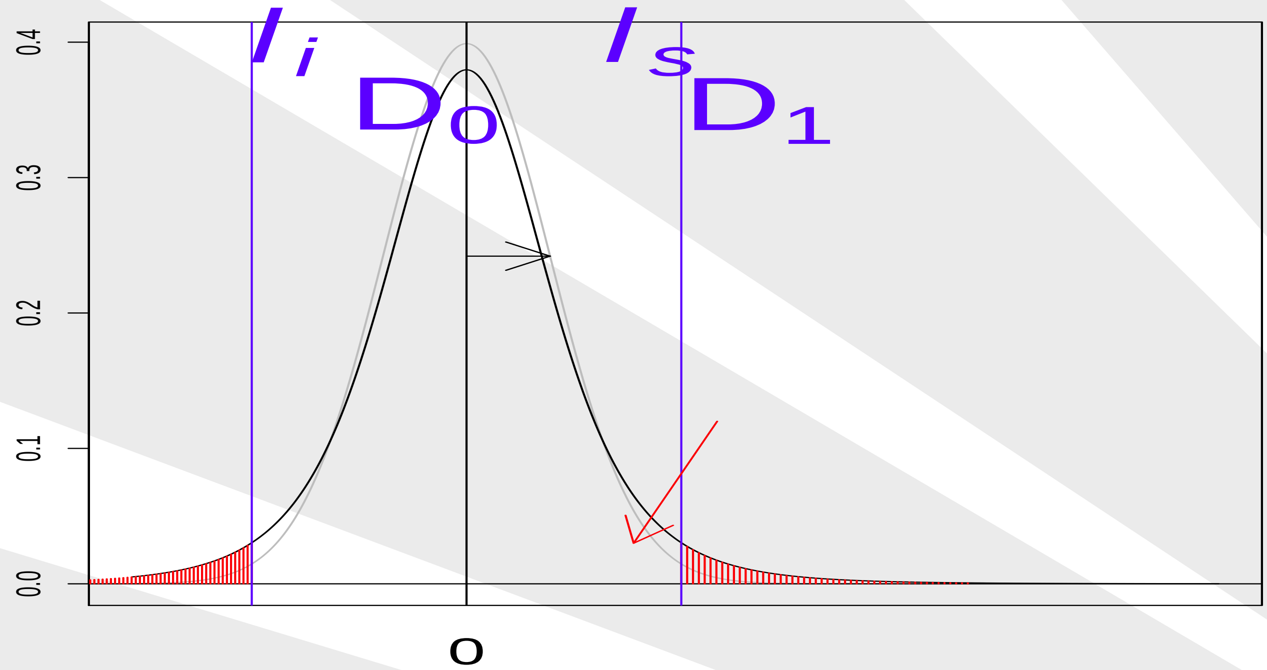
➤ divisée par l'estimateur de son écart-type

➤ on s'intéresse donc à la loi de

$$\frac{\bar{X}-\mu_0}{\frac{S}{\sqrt{n}}}$$

- cette quantité suit une loi de Student à $n-1$ ddl

Limite de décision pour une loi de Student



- pour mémoire on représente la gaussienne (en gris)

Loi dans l’alternative

- lorsque l’on s’intéresse à la loi de la statistique de test sous H_1

➤ on doit considérer la loi de

$$\frac{\bar{X}-\mu_0}{\frac{S}{\sqrt{n}}}$$

- quand X est centrée sur μ_1, cette v.a. n’est plus centrée

$$\begin{aligned}\frac{\bar{X}-\mu_0}{\frac{S}{\sqrt{n}}} &= \frac{\bar{X}-\mu_1+\mu_1-\mu_0}{\frac{S}{\sqrt{n}}}\\ &= \frac{\bar{X}-\mu_1}{\frac{S}{\sqrt{n}}}+\frac{\mu_1-\mu_0}{\frac{S}{\sqrt{n}}}\end{aligned}$$

Loi de Student décentrée

- la quantité

$$\frac{\bar{X}-\mu_1}{\frac{S}{\sqrt{n}}} + \frac{\mu_1-\mu_0}{\frac{S}{\sqrt{n}}}$$

➤ se décompose en deux termes

— $\frac{\bar{X}-\mu_1}{\frac{S}{\sqrt{n}}}$ est une v.a. centrée qui suit une loi de Student

— $\frac{\mu_1-\mu_0}{\frac{S}{\sqrt{n}}}$ qui est une v.a. de décentrage appelée δ

- la v.a. précédente est tabulée et est appelée v.a. de Student décentrée de décentrage $\frac{\mu_1-\mu_0}{\frac{S}{\sqrt{n}}}$

★ cette v.a. permet de calculer la puissance d'un test de Student.

Calcul de la puissance d'un test de Student

- on peut appliquer le résultat précédent au calcul de la puissance en reprenant un exemple proche du précédent.

➤ la limite de la région de décision est

➤ les limites des régions de décision sont pour la gaussienne 1.96 et pour la Student 2.093

➤ les puissances correspondantes sont pour la gaussienne 0.8757 et pour la Student 0.8396

Comparaison d’espérances (de moyennes)

Comparaison d'espérances : modèle

- on reste dans le cas d'une variable d'étude postulée gaussienne
  - ➢ le modèle est donc encore gaussien (comme jusqu'ici)
- et l'on suppose qu'il existe deux traitements différents qui agissent sur l'espérance de la variable
  - ➢ exemple : le rendement pour deux variétés différentes de blé
  - ➢ exemple : le poids pour deux traitements amaigrissants

Comparaison d'espérances : hypothèses

- les hypothèses ne consistent plus à comparer une espérance à une référence

★ mais à comparer deux espérances entre elles

— $H_O : \mu_1 = \mu_2$: il n'y a pas d'effet du traitement (de la variété)

— $H_O : \mu_1 \neq \mu_2$: il y a un effet du traitement (de la variété)

➤ on peut réécrire ces deux hypothèses sous une forme différente qui les rapproche du cas déjà étudié

— $H_O : \mu_1 - \mu_2 = 0$

— $H_O : \mu_1 - \mu_2 \neq 0$

Comparaison d'espérances : statistique de test

- la statistique de test est (naturellement) un estimateur de $\mu_1 - \mu_2$

➤ on peut proposer $\bar{X}_1 - \bar{X}_2$ (moyenne du groupe 1 - moyenne du groupe 2)

- si les variances sont les mêmes pour les deux traitements

➤ la loi de cet estimateur est assez simple

➤ il s'agit d'une variable aléatoire gaussienne

— d'espérance $\mu_1 - \mu_2$
— de variance $\frac{1}{n_1} + \frac{1}{n_2}$

- si la variance σ^2 est connue, on est conduit à un test gaussien

- si la variance σ^2 est inconnue, on est conduit à un test de Student

Cas du test de Student

- la seule difficulté restante est l'estimation de σ^2 : variance commune aux deux traitements
- il existe deux estimateurs potentiels de σ^2

— l'estimateur construit sur le premier échantillon

$$\frac{1}{n_1 - 1} \sum_{r=1}^{n_1} (X_{1r} - \bar{X}_1)^2$$

— l'estimateur construit sur le second échantillon

$$\frac{1}{n_2 - 1} \sum_{r=1}^{n_2} (X_{2r} - \bar{X}_2)^2$$

- il est naturel de chercher à combiner ces deux estimateurs

Estimation optimale de la variance

➤ la combinaison la plus pertinente (et optimale) est la moyenne pondérée (par les degrés de liberté) des deux estimateurs précédents

$$\frac{(n_1-1)\frac{1}{n_1-1}\sum_{r=1}^{n_1}(X_{1r}-\bar{X}_1)^2+(n_2-1)\frac{1}{n_2-1}\sum_{r=1}^{n_2}(X_{2r}-\bar{X}_2)^2}{n_1+n_2-2}$$

Estimateur de σ^2

L'estimateur combiné de la variance pour le test de Student de comparaison de variance est

$$\frac{\sum_{r=1}^{n_1}(X_{1r}-\bar{X}_1)^2+\sum_{r=1}^{n_2}(X_{2r}-\bar{X}_2)^2}{n_1+n_2-2}$$

Comparaison d'espérances : statistique de test

- la construction des régions de décision est classique
  - ➤ région de rejet bilatérale extérieure
- la normalisation à peine plus compliquée (voir TD)
  - ➤ il suffit d'utiliser la statistique de Student