

Probabilités et Statistiques

Année 2009/2010

laurent.carraro@telecom-st-etienne.fr

olivier.roustant@emse.fr

Cours n°12

Théorie des tests statistiques

Test ?

Problème de décision

... en contexte incertain

Exemples :

- Le médicament MEDOC est-il efficace ?
- La machine PROD est-elle bien réglée ?
- Les OGM sont-ils dangereux ?
- L'augmentation de 2% de nos ventes ce dernier mois est-elle significative ?

Points communs aux exemples

- La décision ne peut être certaine ;
- elle sera prise sur la base d'observations ;
- tous les facteurs influents ne sont pas connus, et encore moins mesurés.

A large, thick orange arrow that starts from the bottom left corner of the box containing the list and points towards the text "Utilisation du formalisme probabiliste".

Utilisation du formalisme probabiliste

Vous avez dit hypothèse ?

➤ **On oppose deux hypothèses :**

- MEDOC : efficace vs non efficace
- PROD : bien réglée vs dérégulée
- OGM : dangereux vs inoffensifs

➤ **Notations :**

- H_0 : hypothèse nulle
- H_1 : hypothèse alternative

Qui est H_0 ?

➤ Les deux hypothèses n'ont pas le même rôle

- MEDOC :

- le fabricant pense que le médicament est efficace

➡ H_0 : efficace

- les autorités de santé veulent des preuves

➡ H_0 : inefficace

- OGM ?

- PROD ?

Démarche

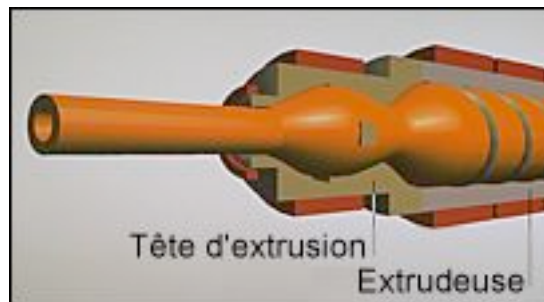
1. On fixe H_0 et H_1 .
2. On évalue une quantité, appelée **score** ou **statistique de test**.
3. Si cette quantité dépasse un certain seuil, on rejette H_0 .
4. On probabilise notre décision...



Un exemple simpl(ist)e

➤ Exemple de type PROD

- Usine de fabrication de tubes pour cosmétiques
- Procédé par extrusion de polymère, puis coupure

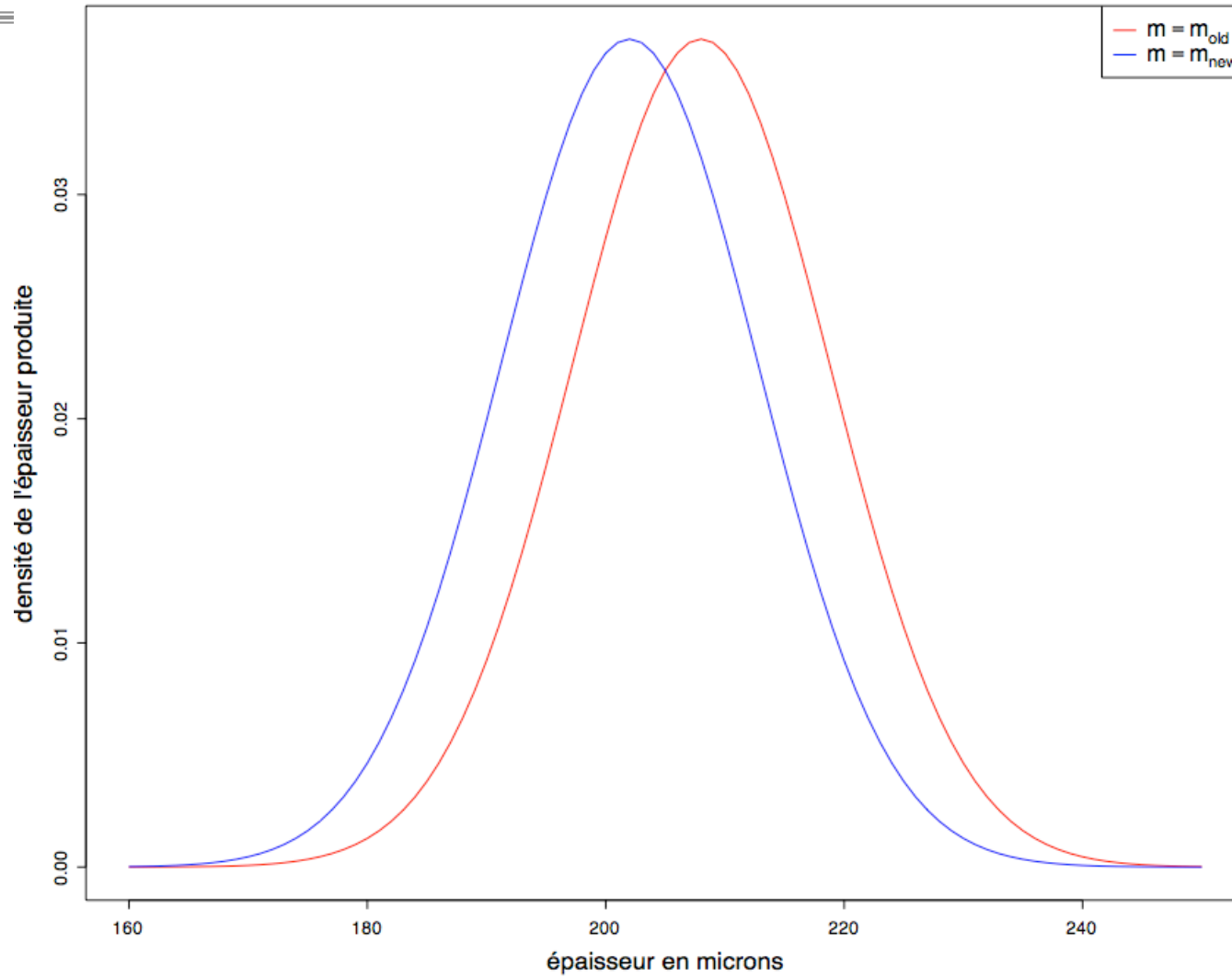


- Paramètre sensible : épaisseur du tube en μm

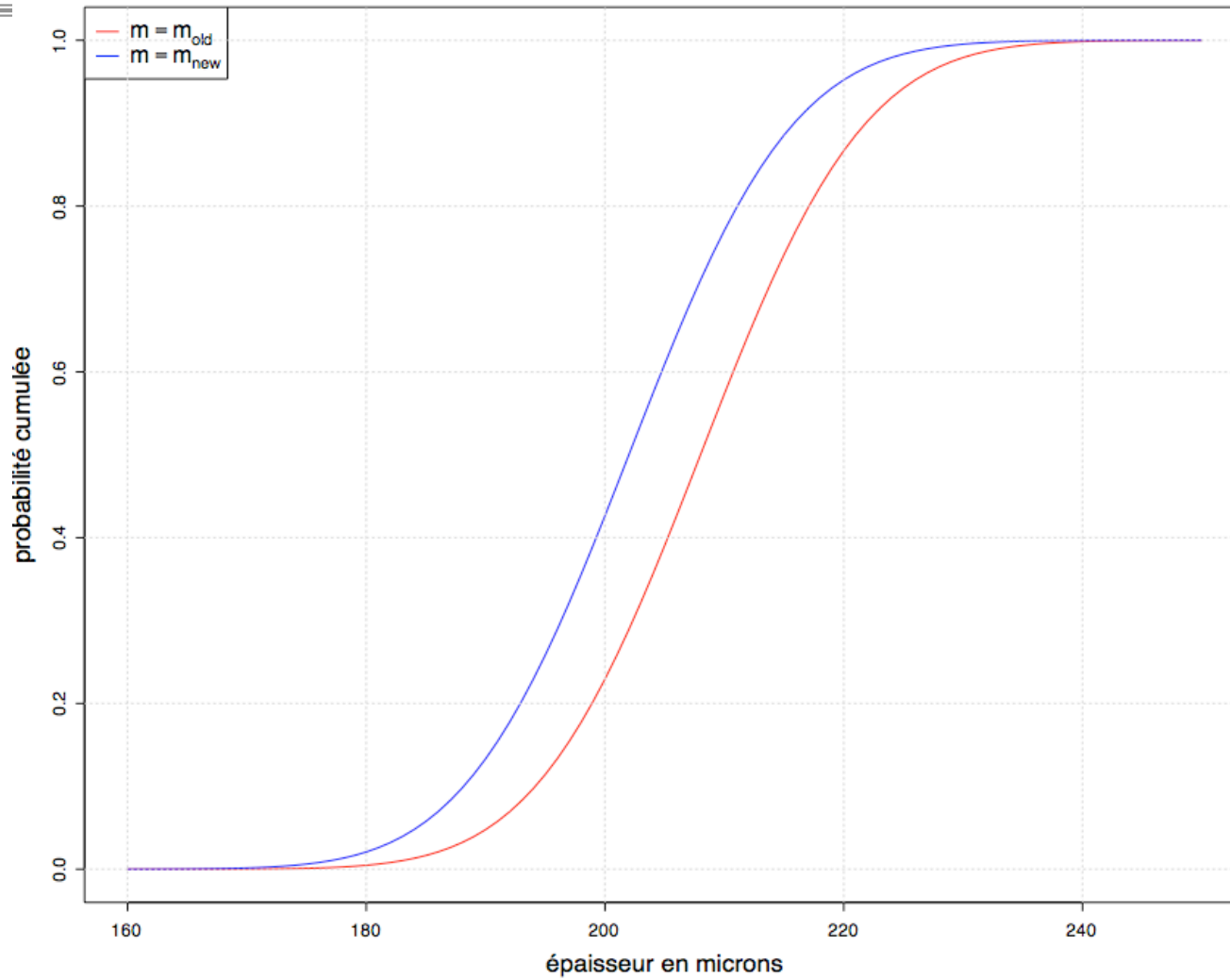
Problème et hypothèses

- En fonctionnement normal, l'épaisseur mesurée d'un tube suit une loi normale $N(m_{\text{old}}, s_{\text{old}}^2)$, où :
 - $m_{\text{old}} = 208 \mu\text{m}$
 - $s_{\text{old}} = 10,8 \mu\text{m}$
- Un changement de fournisseur fait suspecter une diminution de la moyenne : $m_{\text{new}} = 202 \mu\text{m}$.
- On observe 20 épaisseurs de tubes, réalisations indépendantes d'une v.a. de loi normale $N(m, s_{\text{old}}^2)$.
- A-t-on $m = m_{\text{old}}$ ou $m = m_{\text{new}}$?

Densités des productions sous les deux hypothèses



Fonctions de répartition des productions sous les deux hypothèses

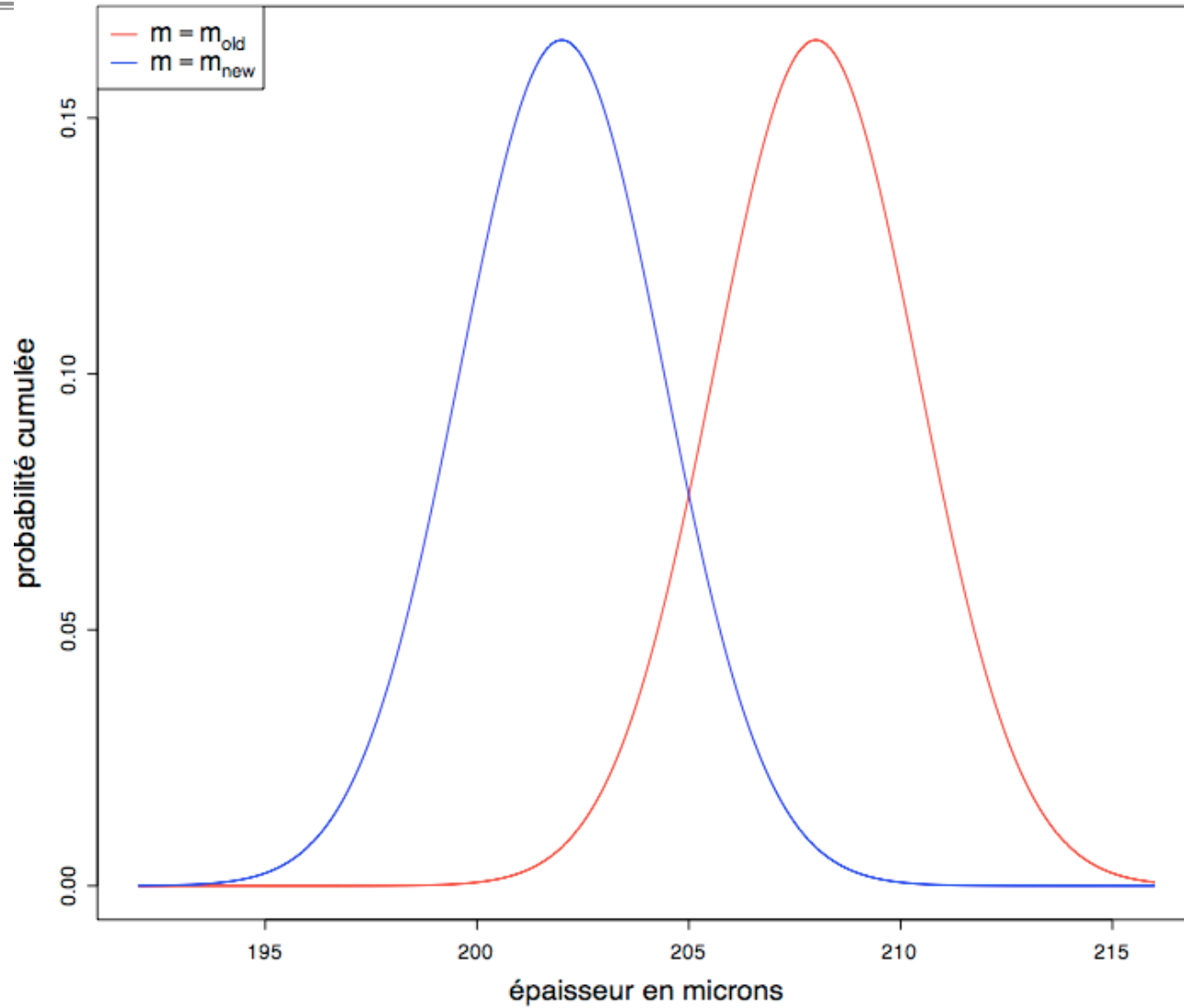


Démarche

- $H_0 : m = m_{\text{new}}$
- Score = épaisseur moyenne \bar{e}
- Décision : si $\bar{e} > \text{seuil}$, on rejette H_0
- On probabilise :
Sous H_0 , \bar{e} est de loi normale $N(m_{\text{new}}, s_{\text{old}}^2/20)$

$$P(\bar{e} > \text{seuil} / H_0) = 1 - \Phi\left(\frac{\text{seuil} - m_{\text{new}}}{s_{\text{old}}/\sqrt{20}}\right)$$

densités de l'épaisseur moyenne sous H0 et H1



Le risque α

- On fixe un niveau de risque α : $\alpha = 5\%$
- On évalue *seuil* pour que :

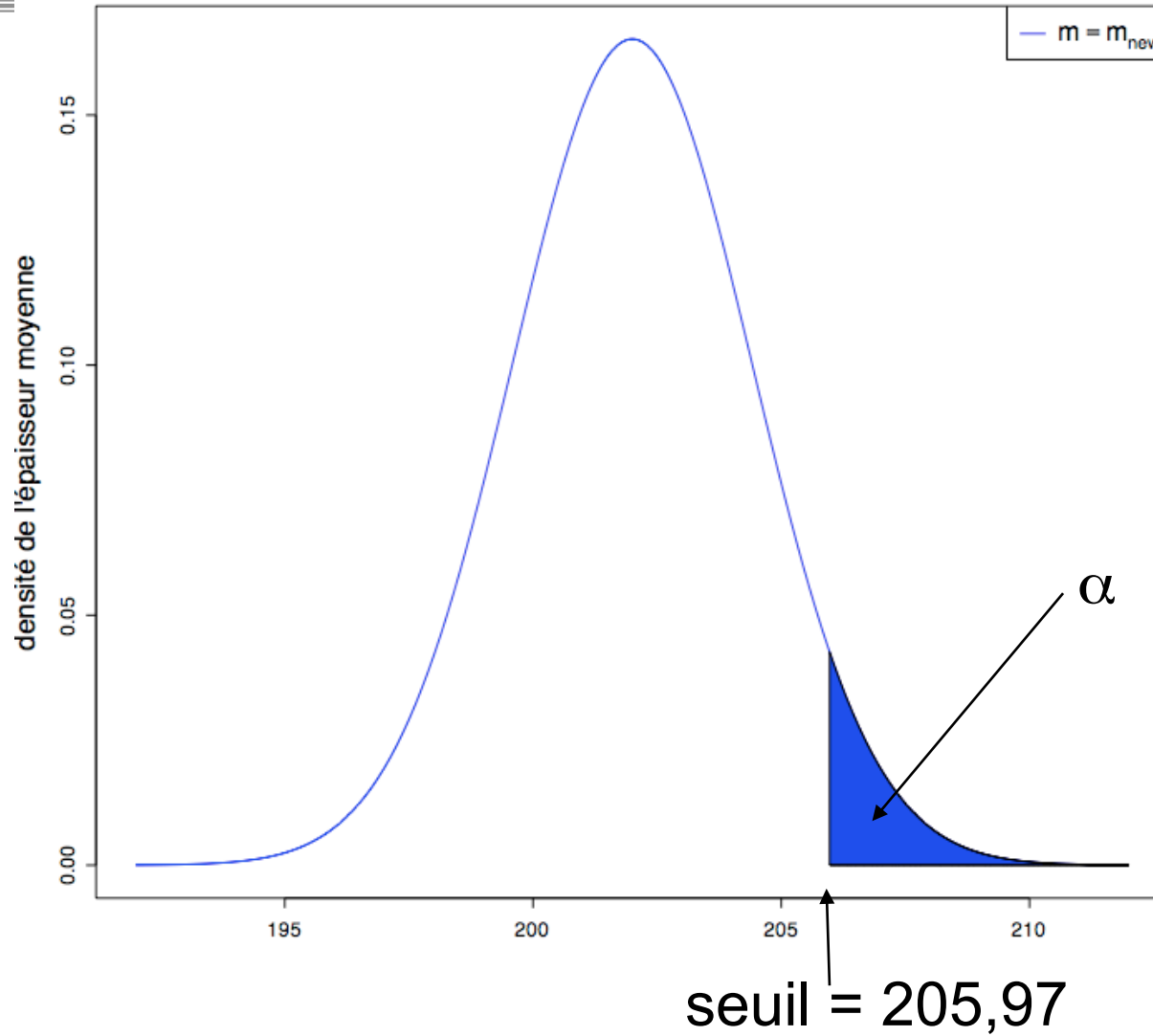
$$P(\bar{e} > \textit{seuil} / H_0) = \alpha$$

Ici, $\textit{seuil} = m_{\text{new}} + 1.64 s_{\text{old}} / \sqrt{20} = 205,97$

- La région $\{\bar{e} > \textit{seuil}\}$ est la région critique.
- Signification ?

Toujours la loi des grands nombres (simulation)

Le risque alpha



Décisions selon les cas

➤ **Supposons :**

1. $\bar{e} = 206,4$
2. $\bar{e} = 207,9$
3. $\bar{e} = 205,2$

➤ **Décisions :**

1. rejet de H_0
2. rejet de H_0
3. on conserve H_0

Le risque β

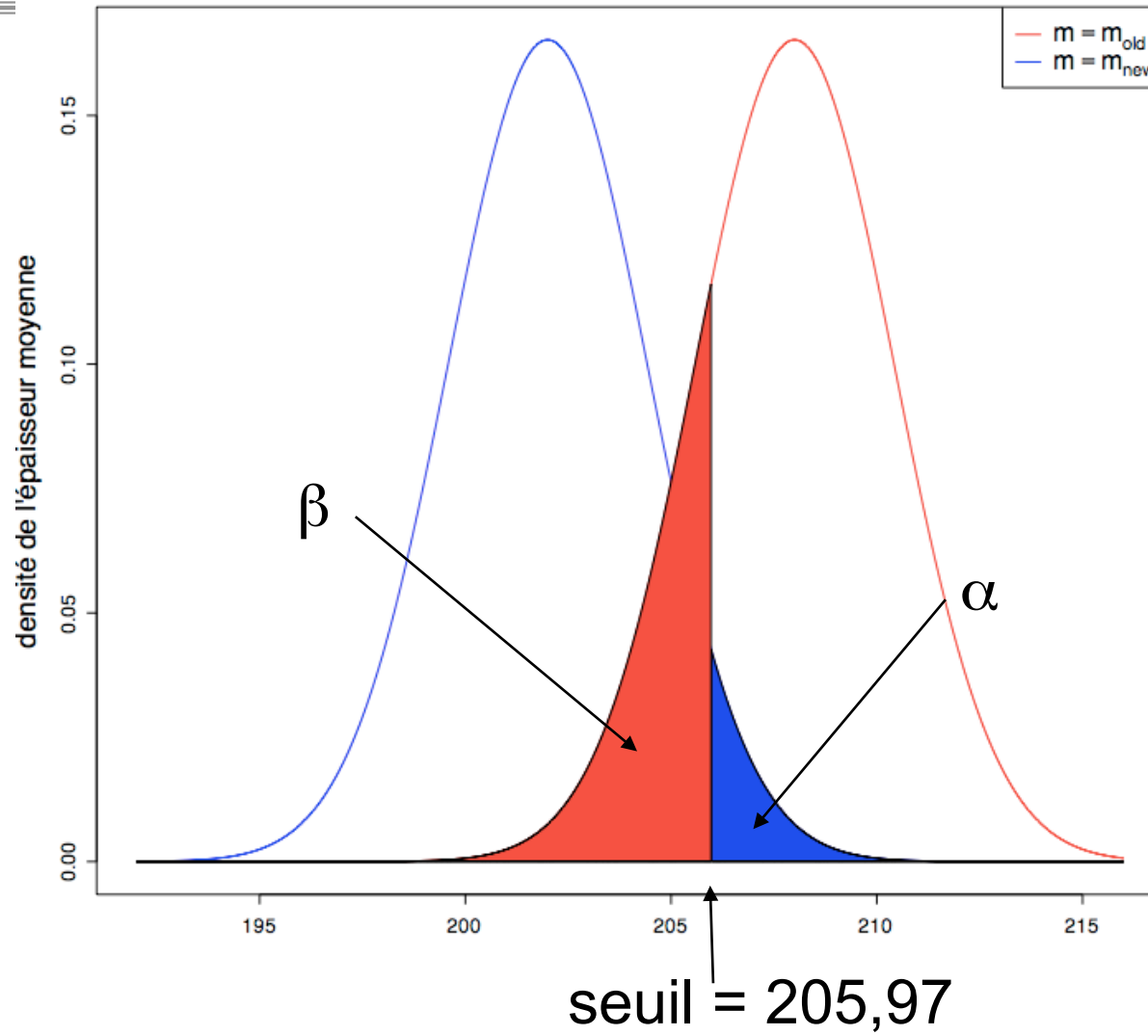
- Si on décide de rejeter H_0 , on a peu de chances de faire erreur (cf. risque α).
- Et si on conserve H_0 , a-t-on raison ??
- Risque de **seconde espèce** β :

$$\beta = P(\bar{e} \leq \text{seuil} / H_1)$$

Ici, $\beta = P(N(202, 10.8^2/20) \leq 205,97) = 20\%$

- α est appelé **risque de première espèce**.

Les risques alpha et bêta



Récapitulons

	Décision		
Réalité		H_0	H_1
H_0		$1-\alpha$	α
H_1		β	$1-\beta$

Déroulement d'un test

1. On fixe H_0 .
2. On définit une région critique (rejet de H_0) à partir d'un score S :
rejet de H_0 si $S \geq \text{seuil}$
3. On fixe α qui détermine seuil tel que :
 $P(S \geq \text{seuil} / H_0) = \alpha$
4. On décide, et si on conserve H_0 , on regarde β

Retour sur le choix de H_0

- Seul α est maîtrisé.
- Exemple PROD :
 - Situation 1 : grosses séries de moyenne qualité :
 - Risque majeur : arrêter la production à tort.
 $\alpha = P(\text{arrêt} / \text{bien réglé}) : H_0 = \text{« bien réglé »}$
 - Situation 2 : CDC client très strict :
 - Risque majeur : produire de mauvais composants.
 $\alpha = P(\text{production} / \text{mal réglé}) : H_0 = \text{« mal réglé »}$

Dernières remarques

- α et β varient en sens contraire.
- Diminution simultanée de α et β possible en augmentant la taille de l'échantillon.
- Critiques :
 - Il se peut qu'aucune des deux hypothèses ne soit correcte (risques de 3ème espèce !!)
 - Si on rejette H_0 avec $\alpha = 5\%$, que donnent 4% ? 1% ? ...

Notion de p-valeur

- Test de région critique de la forme :
rejet de H_0 si $S \geq \text{seuil}$
- On observe s_{obs}
- On évalue la probabilité :
$$p = P(S \geq s_{\text{obs}} / H_0)$$
- p est appelée p-valeur (p-value)

Retour sur l'exemple

1. Cas où $\bar{e} = 206,4$:

$$\text{p-valeur} = P(N(202, 10.8^2/20) > 206,4) = 0.034$$

2. Cas où $\bar{e} = 207,9$:

$$\text{p-valeur} = P(N(202, 10.8^2/20) > 207,9) = 0.0073$$

3. Cas où $\bar{e} = 205,2$:

$$\text{p-valeur} = P(N(202, 10.8^2/20) > 205,2) = 0.093$$

p-valeur pour une épaisseur moyenne observée de 205,2

