

Nombre de sujets nécessaires

Dr Julien Mancini

julien.mancini@univmed.fr

*Laboratoire d'Enseignement et de Recherche
sur le Traitement de l'Information Médicale*

Faculté de Médecine de Marseille, Université de la Méditerranée

Plan

1. Rappels
2. Situation
3. Puissance
4. Nombre de sujets nécessaires
5. Calcul
6. Comparaison de 2 pourcentages
7. Conclusion

I. Rappels

1. Rappels

2. Situation
3. Puissance
4. NSN
5. Calcul
7. Pourcentage
6. Conclusion

- Variabilité individuelle:

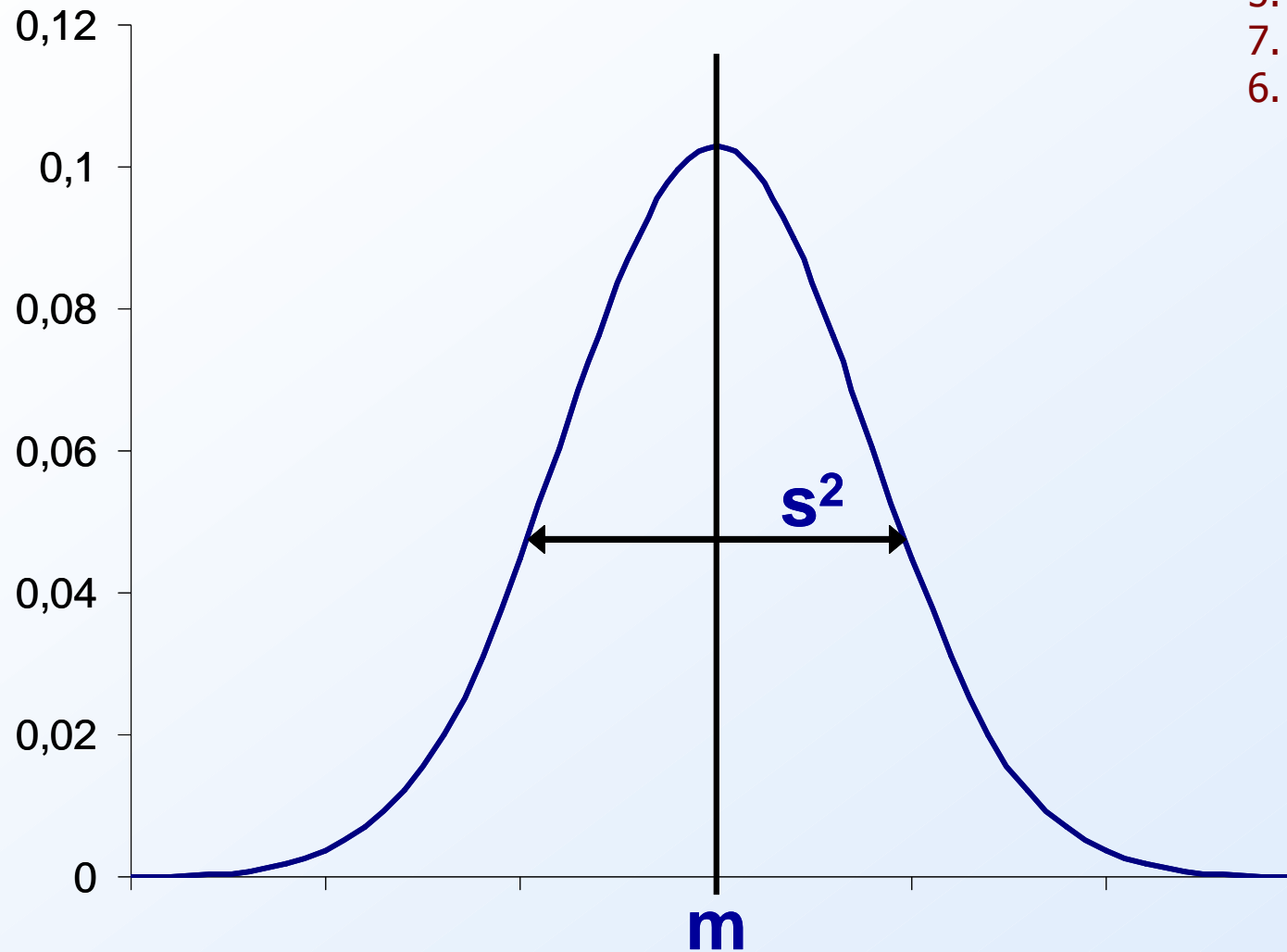
⇒ Individus tous différents

⇒ Variable: mesures différentes /individus

⇒ Distribution de la variable

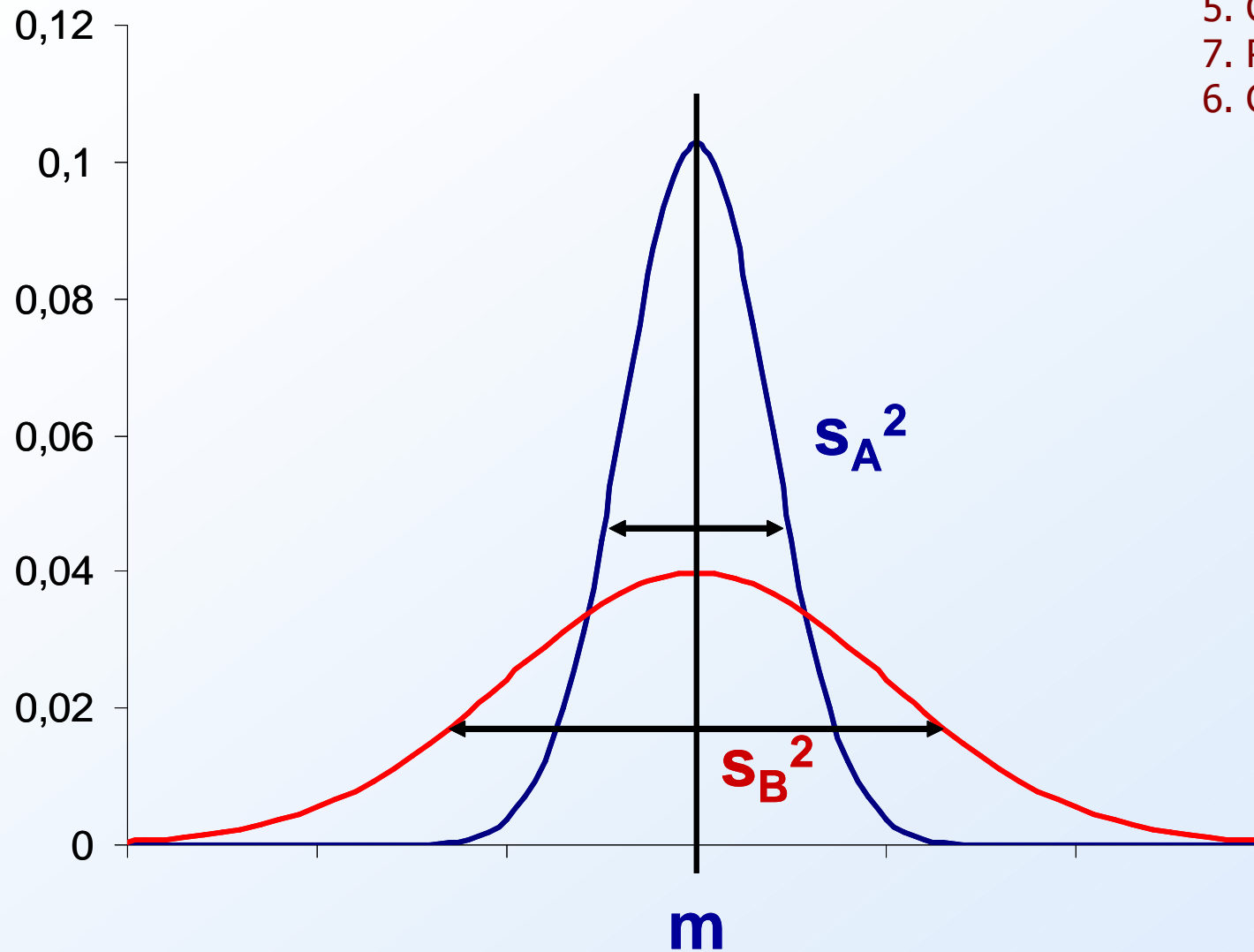
1. Rappels

2. Situation
3. Puissance
4. NSN
5. Calcul
6. Conclusion
7. Pourcentage



1. Rappels

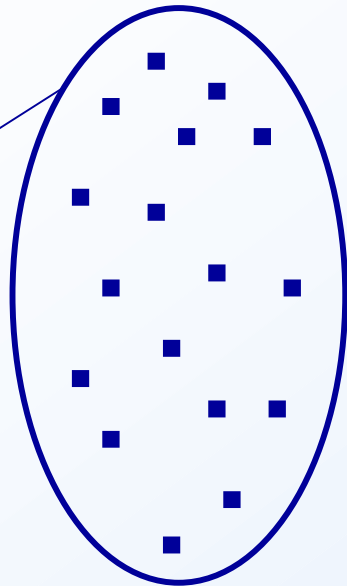
- 2. Situation
- 3. Puissance
- 4. NSN
- 5. Calcul
- 7. Pourcentage
- 6. Conclusion



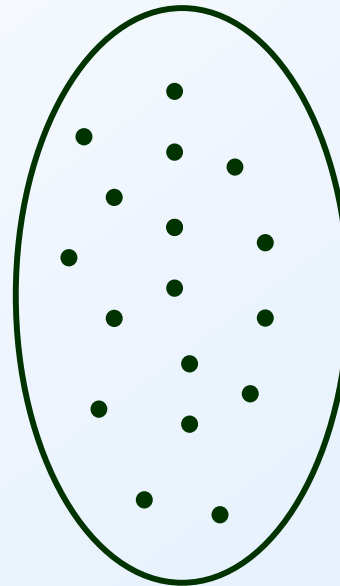
1. Rappels

2. Situation
3. Puissance
4. NSN
5. Calcul
7. Pourcentage
6. Conclusion

Groupe A



=



Groupe B

n_A : nombre de sujets

m_A : moyenne

s^2_A : variance

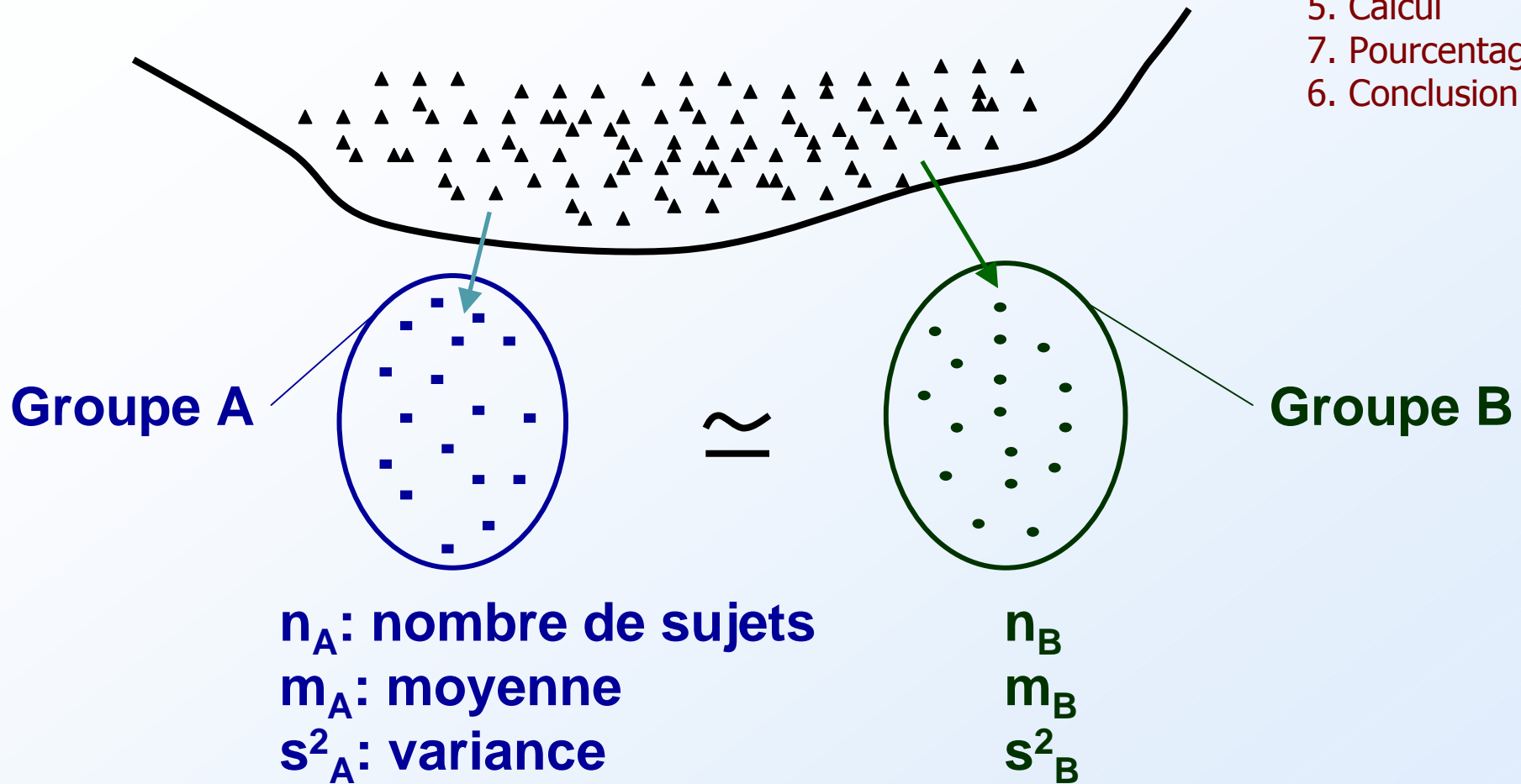
n_B

m_B

s^2_B

1. Rappels

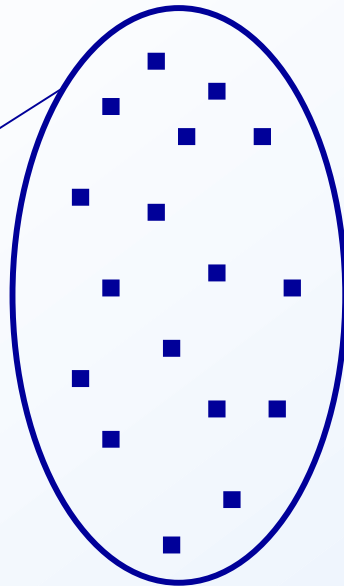
2. Situation
3. Puissance
4. NSN
5. Calcul
7. Pourcentage
6. Conclusion



1. Rappels

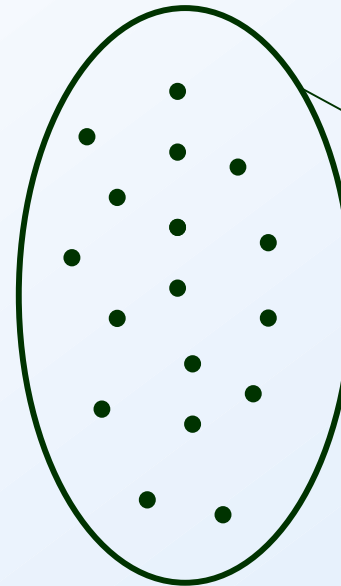
2. Situation
3. Puissance
4. NSN
5. Calcul
7. Pourcentage
6. Conclusion

Groupe A



\neq

Groupe B



n_A : nombre de sujets

m_A : moyenne

s^2_A : variance

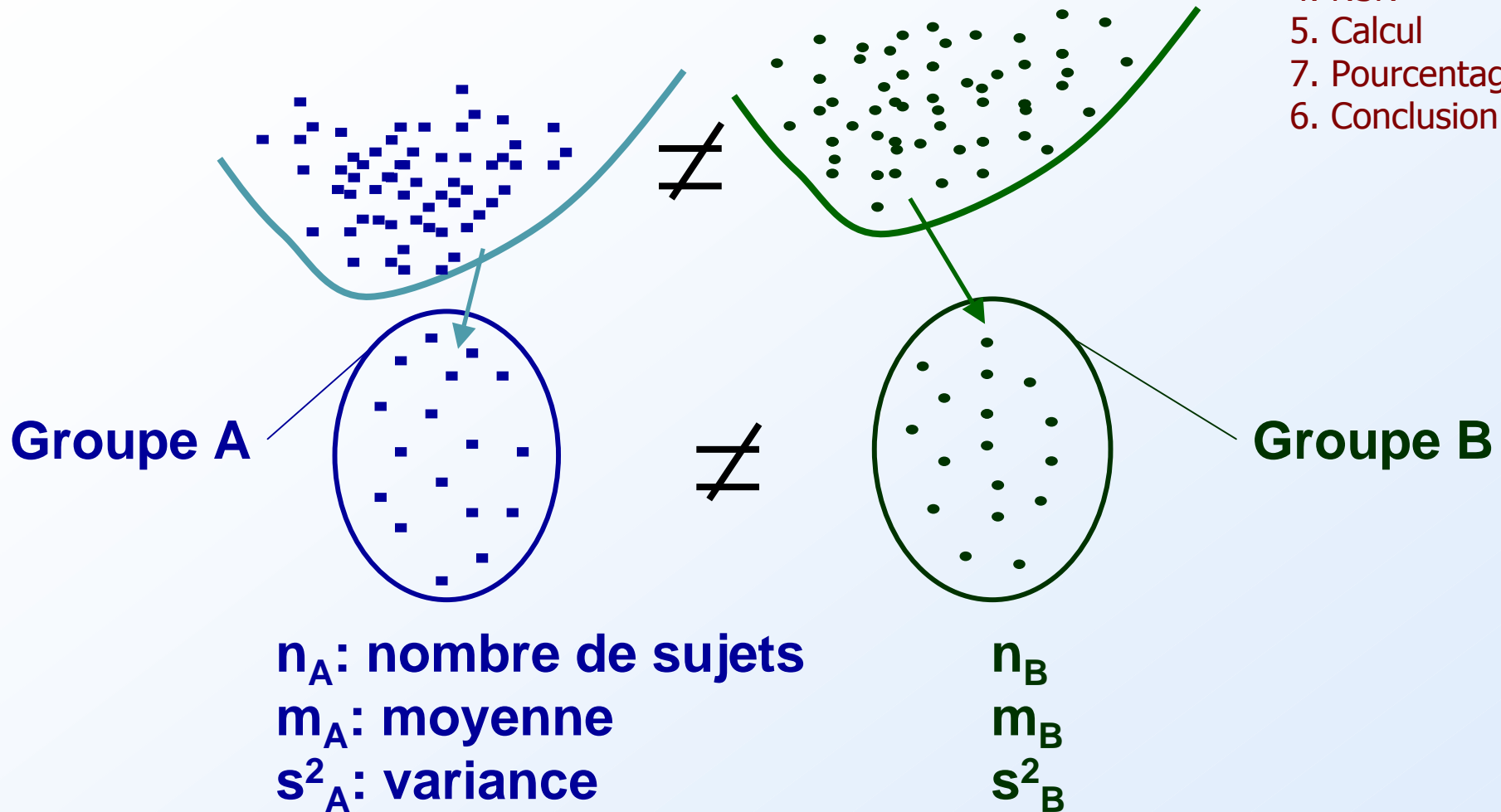
n_B

m_B

s^2_B

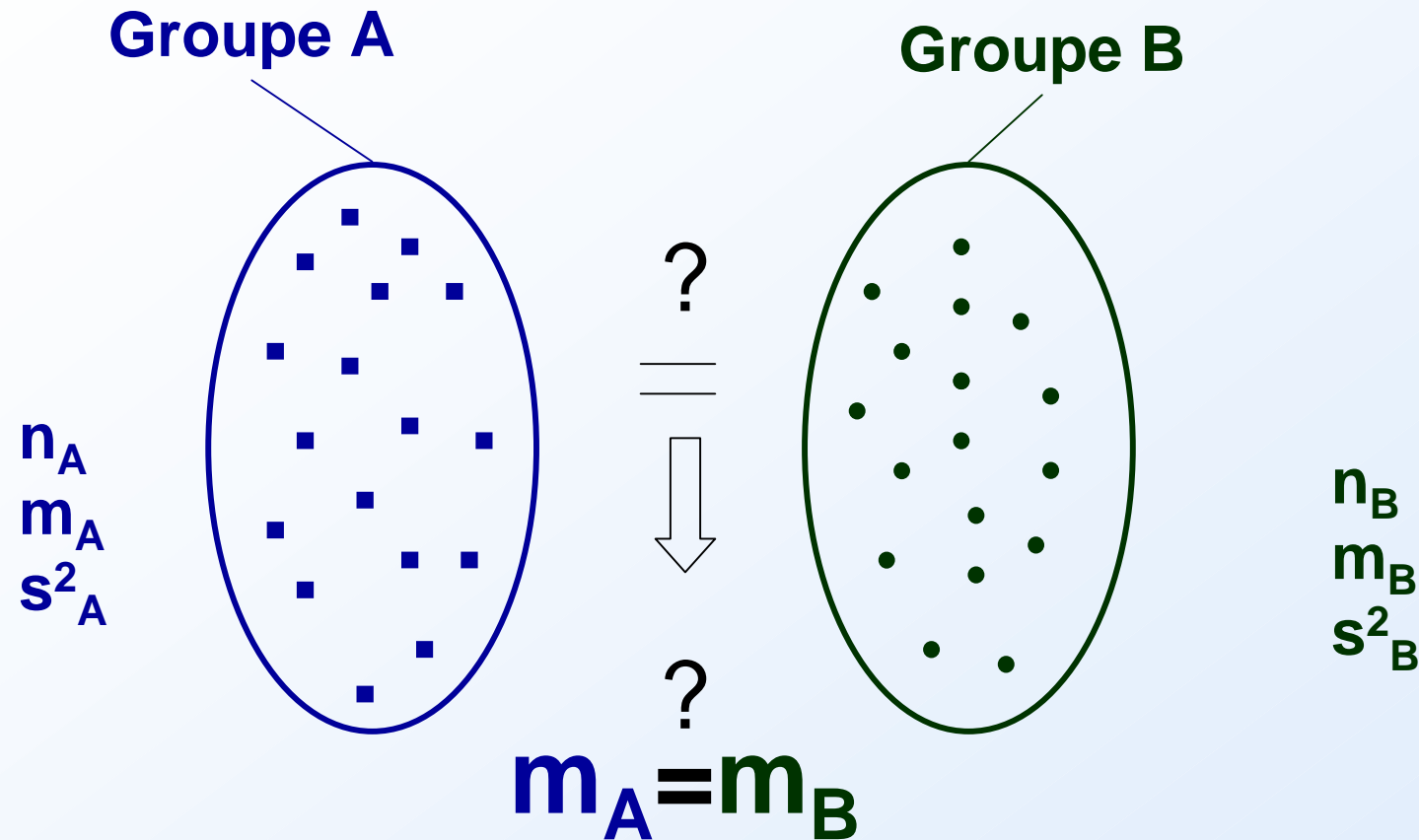
1. Rappels

2. Situation
3. Puissance
4. NSN
5. Calcul
7. Pourcentage
6. Conclusion



1. Rappels

2. Situation
3. Puissance
4. NSN
5. Calcul
7. Pourcentage
6. Conclusion



Test de comparaison de 2 moyennes

II. Situation

1. Rappels
- 2. Situation**
3. Puissance
4. NSN
5. Calcul
6. Pourcentage
7. Conclusion

- Avant de débiter une étude:
⇒ Construction du protocole
- *Exemple:*
 - *Facteur de risque de paludisme : présent (A)/absent (B)*
 - *Parasitémie à P. falciparum chez les enfants*
- Mais
 - 2 groupes ⇒ 2 observations différentes
 - Si la différence existe
⇒ pouvoir séparer les 2 groupes

Pouvoir séparateur

1. Rappels

2. Situation

3. Puissance

4. NSN

5. Calcul

6. Pourcentage

7. Conclusion

- Analogie: le microscope

Si le biologiste ne voit rien

⇒ augmenter le grossissement

⇒ refaire le prélèvement



"ne rien voir" \neq "n'existe pas"

Pouvoir Séparateur = Puissance

1. Rappels

2. **Situation**

3. Puissance

4. NSN

5. Calcul

6. Pourcentage

7. Conclusion

- Analogie: le microscope

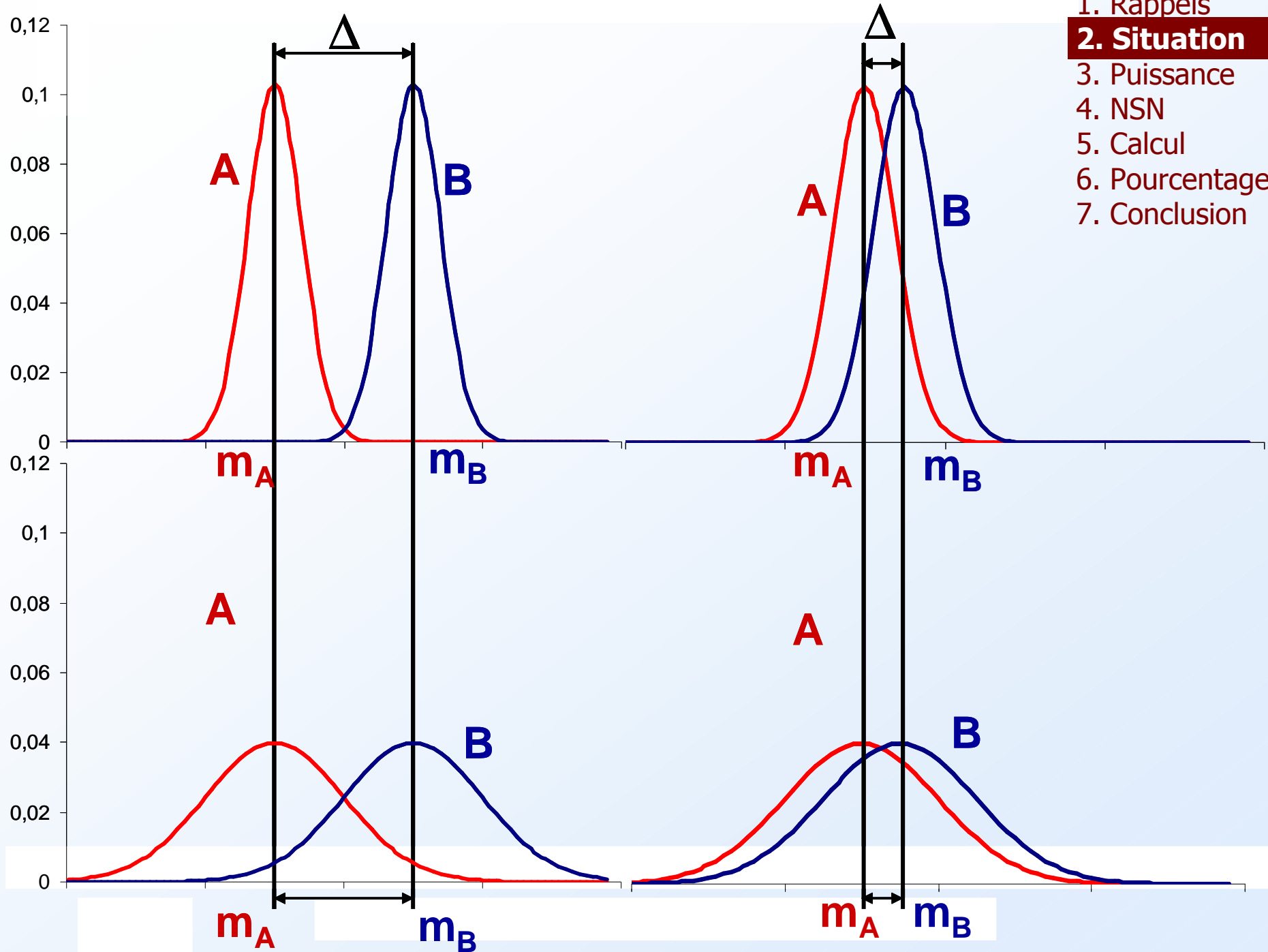
- Augmenter le grossissement :
dépend de la **dimension**

- ⇒ Pour un **test** : ⇒ dépend de **l'effet $m_A - m_B$**
 - ⇒ prendre plus **d'enfants**

- Refaire le prélèvement :

- ⇒ Pour un **test** : prendre un autre **échantillon**
(biais de sélection ?)

- 1. Rappels
- 2. Situation**
- 3. Puissance
- 4. NSN
- 5. Calcul
- 6. Pourcentage
- 7. Conclusion



III. Puissance

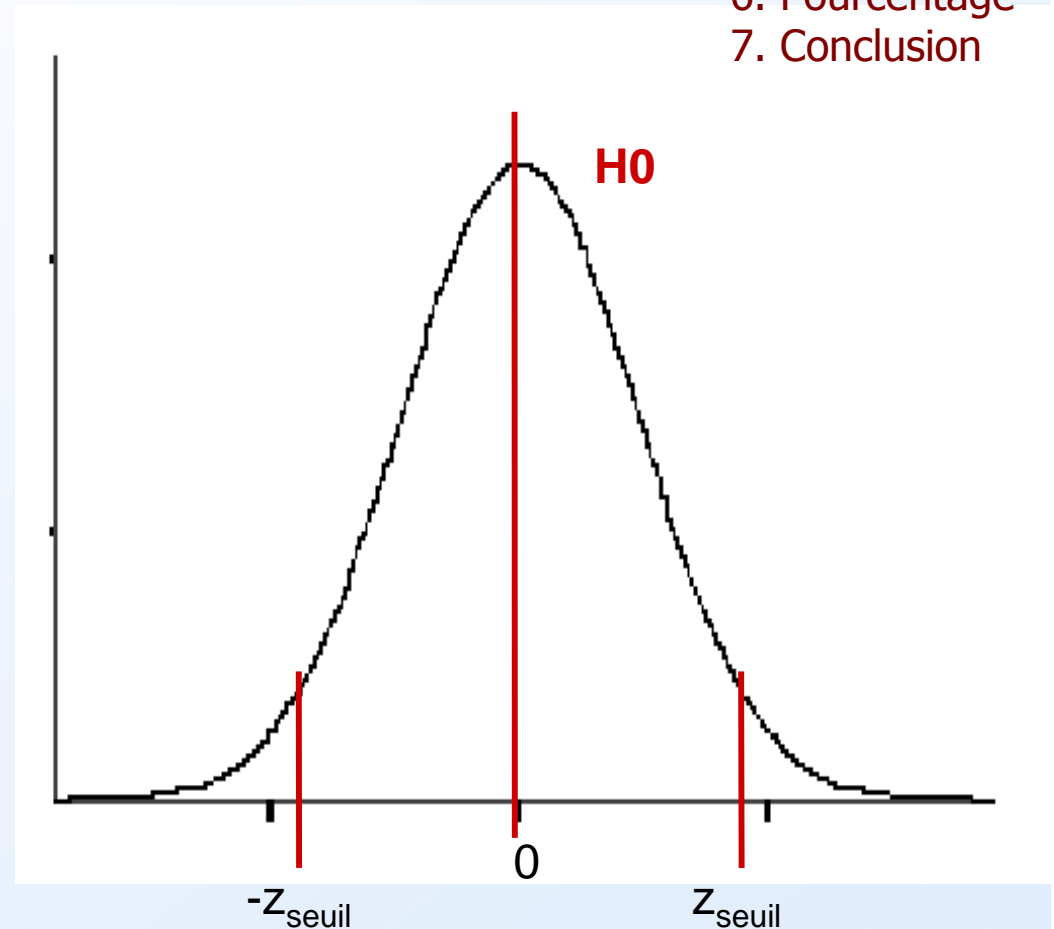
1. Rappels
2. Situation
- 3. Puissance**
4. NSN
5. Calcul
6. Pourcentage
7. Conclusion

2 Risques

Risque de 1ère espèce

$\alpha = \text{prob}(\text{rejet } H_0 / H_0 \text{ vraie})$

$= \text{prob}(|Z| \geq z_{\text{seuil}} / H_0 \text{ vraie})$



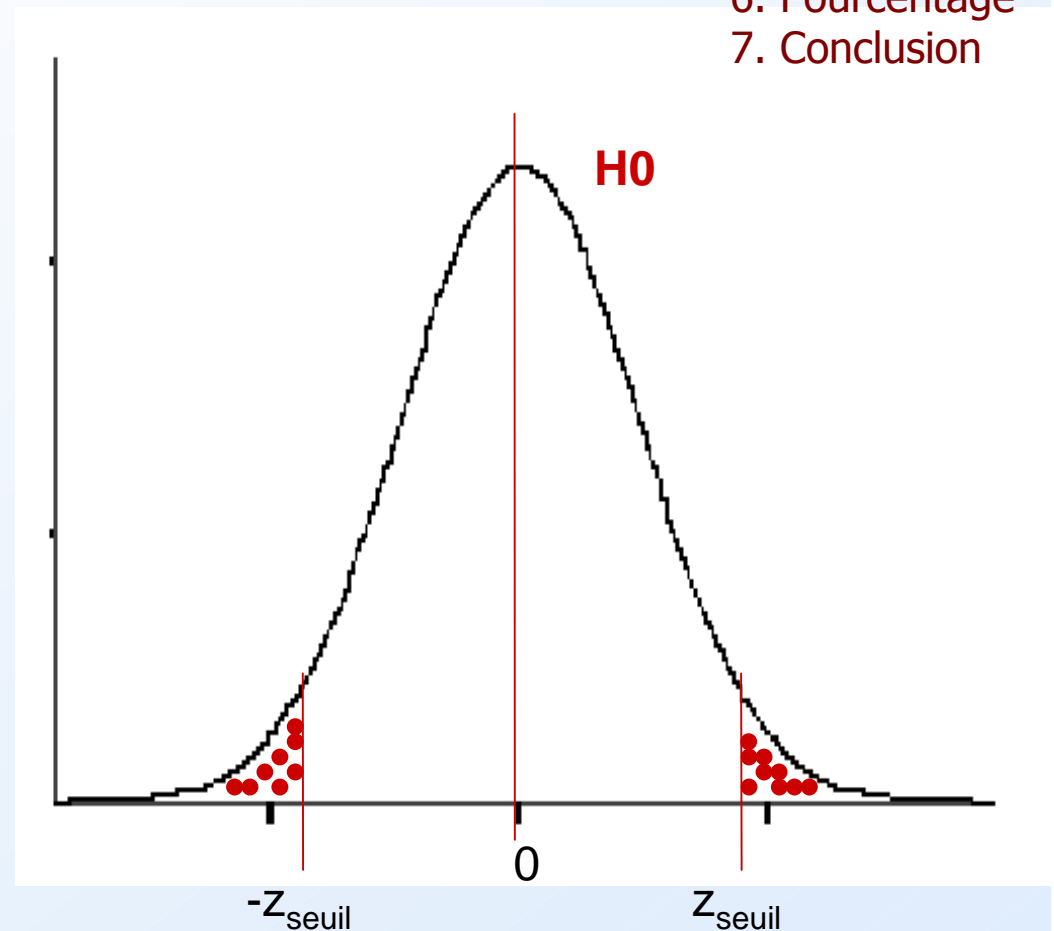
1. Rappels
2. Situation
- 3. Puissance**
4. NSN
5. Calcul
6. Pourcentage
7. Conclusion

2 Risques

Risque de 1ère espèce

$\alpha = \text{prob}(\text{rejet } H_0 / H_0 \text{ vraie})$

$= \text{prob}(|Z| \geq z_{\text{seuil}} / H_0 \text{ vraie})$



1. Rappels
2. Situation
- 3. Puissance**
4. NSN
5. Calcul
6. Pourcentage
7. Conclusion

Risque de 2ème espèce

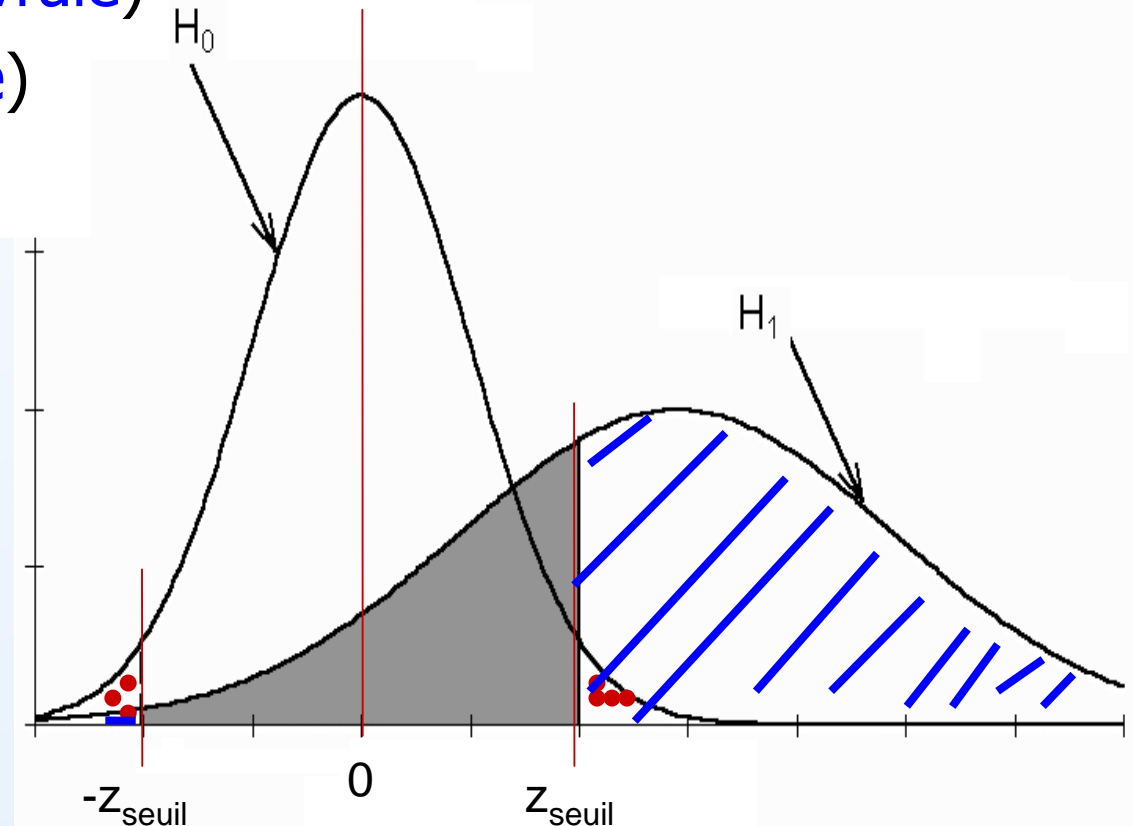
$$\beta = \text{prob}(\text{non rejet } H_0 / H_1 \text{ vraie})$$
$$= \text{prob}(|Z| < z_{\text{seuil}} / H_1 \text{ vraie})$$

Puissance

$$1 - \beta$$

$$= \text{prob}(\text{rejet } H_0 / H_1 \text{ vraie})$$

$$= \text{prob}(|Z| \geq z_{\text{seuil}} / H_1 \text{ vraie})$$



Puissance

1. Rappels
2. Situation
- 3. Puissance**
4. NSN
5. Calcul
6. Pourcentage
7. Conclusion

Capacité d'un test à montrer une différence

Dépend:

de la différence minimale d'intérêt Δ

du **nombre de sujets**

de la variance

du risque α

IV. NSN

1. Rappels
2. Situation
3. Puissance
- 4. NSN**
5. Calcul
6. Pourcentage
7. Conclusion

Objectif :

Pouvoir séparer les 2 groupes

⇒ **Combien de sujets** faut-il inclure dans chaque groupe ?

Définition

1. Rappels
2. Situation
3. Puissance
- 4. NSN**
5. Calcul
6. Pourcentage
7. Conclusion

Nombre de sujets nécessaire

- ⇒ **pouvoir séparer** 2 groupes (=puissance)
- ⇒ pour un **effet** (=différence) donné
- ⇒ pour une **variance** donnée
- ⇒ avec un **risque d'erreur** fixé

Paramètres

1. Rappels
2. Situation
3. Puissance
- 4. NSN**
5. Calcul
6. Pourcentage
7. Conclusion

1. Le pouvoir séparateur

- Puissance d'un test statistique:

**Capacité à montrer
un effet
lorsqu'il existe**

- En général puissance \geq **80%**

2. L'effet

- **Différence minimale d'intérêt:**

$$\Delta = m_A - m_B$$

Exemple :

*Facteur de risque **présent (A)** vs **absent (B)***

Groupe A: **m_A** = 5000 parasites / μ l

Groupe B: **m_B** = 4500 parasites / μ l

} Pas d'intérêt

Groupe A: **m_A** = 5000 parasites / μ l

Groupe B: **m_B** = 600 parasites / μ l

} Intérêt++

L'effet

1. Rappels
2. Situation
3. Puissance
- 4. NSN**
5. Calcul
6. Pourcentage
7. Conclusion

- Choix difficile
- Fonction du problème
- Critères:
 - Biologiques
 - Cliniques

 - *Pas statistique*

3. La variance

- Variabilité de l'ensemble

Donnée par la connaissance:

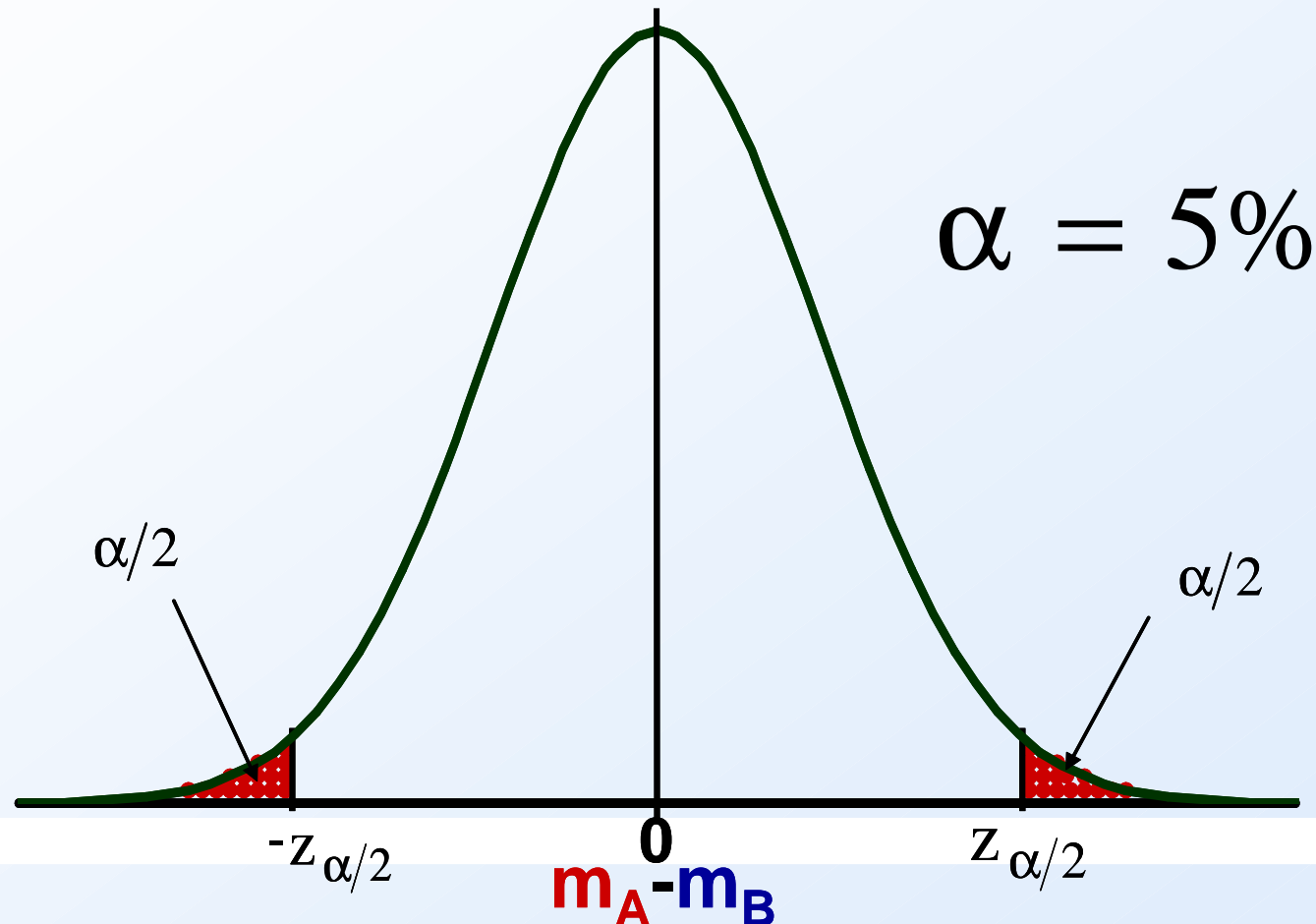
- Littérature
- Étude préliminaire

Paramètres

1. Rappels
2. Situation
3. Puissance
- 4. NSN**
5. Calcul
6. Pourcentage
7. Conclusion

4. Le risque d'erreur

- Séparer 2 groupes **à tort**



V. Calcul

1. Rappels
2. Situation
3. Puissance
4. NSN
- 5. Calcul**
6. Pourcentage
7. Conclusion

- Déterminer les 4 paramètres

- Puissance **80%** $\Leftrightarrow z_{\text{puis.}} = -0,842$ (loi Normale)

- Différence minimale d'intérêt **$\Delta = 3000$**

- Variance (étude préliminaire) : **$s^2 = 75.10^6$**

- Risque **$\alpha = 5\%$** $\Leftrightarrow z_{\alpha/2} = 1,96$ (loi Normale)

Formule

1. Rappels
2. Situation
3. Puissance
4. NSN
- 5. Calcul**
6. Pourcentage
7. Conclusion

nombre de sujets nécessaires par groupe :

$$n_A = 2 \times \frac{s^2}{\Delta^2} \times \left(z_{\alpha/2} - z_{\text{puis}} \right)^2$$

Formule

1. Rappels
2. Situation
3. Puissance
4. NSN
- 5. Calcul**
6. Pourcentage
7. Conclusion

nombre de sujets nécessaires par groupe :

The diagram shows the formula $n_A = 2 \times \frac{s^2}{\Delta^2} \times (z_{\alpha/2} - z_{\text{puis}})^2$. The term s^2 is circled in blue, with a blue arrow pointing from a box labeled "connaissance" above it. The term Δ^2 is also circled in blue, with a blue arrow pointing from a box labeled "À définir" below it.

$$n_A = 2 \times \frac{s^2}{\Delta^2} \times (z_{\alpha/2} - z_{\text{puis}})^2$$

Formule

1. Rappels
2. Situation
3. Puissance
4. NSN
- 5. Calcul**
6. Pourcentage
7. Conclusion

nombre de sujets nécessaires par groupe :

$$n_A = 2 \times \frac{s^2}{\Delta^2} \times \left(z_{\alpha/2} - z_{\text{puis}} \right)^2$$

Diagram annotations:

- A blue box labeled "connaissance" points to the s^2 term in the numerator.
- A blue box labeled "À définir" points to the Δ^2 term in the denominator.
- A red box labeled " $\alpha=5\%$
 $\Leftrightarrow 1,96$ " points to the $z_{\alpha/2}$ term in the parentheses.

Formule

1. Rappels
2. Situation
3. Puissance
4. NSN
- 5. Calcul**
6. Pourcentage
7. Conclusion

nombre de sujets nécessaires par groupe :

$$n_A = 2 \times \frac{s^2}{\Delta^2} \times (z_{\alpha/2} - z_{\text{puis}})^2$$

Diagram illustrating the formula for the number of subjects per group (n_A).

- s^2 (variance) is linked to "connaissance" (knowledge).
- Δ^2 (difference) is linked to "À définir" (to be defined).
- $z_{\alpha/2}$ (critical value) is linked to $\alpha=5\% \Leftrightarrow 1,96$.
- z_{puis} (power) is linked to $1-\beta=80\% \Leftrightarrow -0,842$.

Calcul

1. Rappels
2. Situation
3. Puissance
4. NSN
- 5. Calcul**
6. Pourcentage
7. Conclusion

nombre d'enfants nécessaires par groupe :

$$n_A = 2 \times \frac{75 \cdot 10^6}{3000^2} \times (1,96 - (-0,842))^2$$

$$n_A = 130,85$$

Résultat

1. Rappels
2. Situation
3. Puissance
4. NSN
- 5. Calcul**
6. Pourcentage
7. Conclusion

Il faut inclure **131** enfants par groupe pour que:
on ait **80%** de **chance**
de détecter un **effet** de **+3000**
pour une **variance** de **$s^2=75.10^6$**
et avec un **risque** de **$\alpha=5\%$**

Application

1. Rappels
2. Situation
3. Puissance
4. NSN
- 5. Calcul**
6. Pourcentage
7. Conclusion

Conditions d'applications de la formule

$$n_A \text{ et } n_B \geq 30$$



ou

Distributions Normales

Inclusion

1. Rappels
2. Situation
3. Puissance
4. NSN
- 5. Calcul**
6. Pourcentage
7. Conclusion

Nombre d'enfants à inclure :

➤ Au cours de l'étude : **10%** de perte
⇒ inclure +10%

➤ **Au total** : $131+14=$ **145** enfants par groupe

VI. Pourcentages

Comparaison de 2 pourcentages observées
(Test Bilatéral)

1. Rappels
2. Situation
3. Puissance
4. NSN
5. Calcul
- 6. Pourcent.**
7. Conclusion

$$n_A = \frac{\left(z_{\alpha/2} - z_{\text{puis}}\right)^2}{2\left(\arcsin \sqrt{P_1} - \arcsin \sqrt{P_2}\right)^2}$$

VI. Pourcentages

1. Rappels
2. Situation
3. Puissance
4. NSN
5. Calcul
- 6. Pourcent.**
7. Conclusion

Comparaison de 2 pourcentages observées
(Test Bilatéral)

$$n_A = \frac{\left(z_{\alpha/2} - z_{\text{puis}} \right)^2}{2 \left(\arcsin \sqrt{P_1} - \arcsin \sqrt{P_2} \right)^2}$$

Diagram illustrating the formula for sample size n_A in a bilateral test comparing two percentages. The terms P_1 and P_2 are circled in blue. A box labeled "connaissance" (knowledge) points to P_1 , and a box labeled "À définir" (to be defined) points to P_2 .

VI. Pourcentages

1. Rappels
2. Situation
3. Puissance
4. NSN
5. Calcul
- 6. Pourcent.**
7. Conclusion

Comparaison de 2 pourcentages observées
(Test Bilatéral)

$$n_A = \frac{(z_{\alpha/2} - z_{\text{puis}})^2}{2(\arcsin \sqrt{P_1} - \arcsin \sqrt{P_2})^2}$$

$\alpha=5\%$
 $\Leftrightarrow 1,96$

connaissance

À définir

VI. Pourcentages

1. Rappels
2. Situation
3. Puissance
4. NSN
5. Calcul
- 6. Pourcent.**
7. Conclusion

Comparaison de 2 pourcentages observées
(Test Bilatéral)

$$n_A = \frac{(z_{\alpha/2} - z_{\text{puis}})^2}{2(\arcsin \sqrt{P_1} - \arcsin \sqrt{P_2})^2}$$

Diagram illustrating the components of the sample size formula n_A for a bilateral test comparing two percentages:

- $\alpha = 5\%$ (Significance level) is associated with $z_{\alpha/2} = 1,96$.
- $1 - \beta = 80\%$ (Power) is associated with $z_{\text{puis}} = -0,842$.
- P_1 (Percentage 1) is associated with the label "connaissance" (knowledge).
- P_2 (Percentage 2) is associated with the label "À définir" (to be defined).

Exemple

1. Rappels
2. Situation
3. Puissance
4. NSN
5. Calcul
- 6. Pourcent.**
7. Conclusion

Nouvelle étude :

- *Complication du paludisme chez les enfants*
- *Facteur de risque de complication :*
 - *présent (A) / absent (B)*
- *Risque de complication :*
 - *Avec le facteur de risque = **20%***
 - *Contre **10%** dans la population de référence*
*(test bilatéral, puissance de **80%** et seuil de significativité fixé à **5%**)*

Calcul

1. Rappels
2. Situation
3. Puissance
4. NSN
5. Calcul
- 6. Pourcent.**
7. Conclusion

$$n_A = \frac{(1,96 - (-0,842))^2}{2(\arcsin \sqrt{0.10} - \arcsin \sqrt{0.20})^2}$$

$$n_A = \frac{(2,802)^2}{2(0,142)^2} = 194,7$$

Résultat

1. Rappels
2. Situation
3. Puissance
4. NSN
5. Calcul
- 6. Pourcent.**
7. Conclusion

Il faut inclure **195** enfants par groupe pour :

- avoir **80%** de **chance**
- de détecter un **effet** de **+10%**
- pour une **probabilité de base** de **10%**
- avec un **risque** de **$\alpha=5%$**

VII. Conclusion

1. Rappels
2. Situation
3. Puissance
4. NSN
5. Calcul
6. Pourcentage
- 7. Conclusion**

Nombre de sujets
nécessaire pour qu'un **test statistique**
puisse
avoir la **puissance** suffisante
pour montrer un **effet** minimum

- Défaut de méthodologie
 - ⇒ ne pas voir l'effet
 - d'un **facteur de risque**
 - d'un **nouveau traitement**

Références

Jean Bouyer: *Méthodes statistiques, Médecine-Biologie*,
éditions INSERM

Contact

julien.mancini@univmed.fr

Labo. d'Enseignement et de Recherche sur le Traitement de
l'Information Médicale

<http://cybertim.timone.univ-mrs.fr/>