GdR MBGS
MATÉRIAUX de CONSTRUCTION
BIO et GÉO SOURCÉS



Statistiques pour expérimentateurs

Comprendre la variabilité, tester des hypothèses et planifier ses mesures

Romain Clerc (GeM, Nantes Université)

Toutes les données utilisées dans ce cours sont simulées.

GeM Institut de Recherche en Génie Civil et Mécanique
Nantes Université
CENTRALE NANTES
cnrs

Trois éprouvettes,
une moyenne, un écart-type :
c’est l’usage, mais que sait-on ?

Exemple pratique

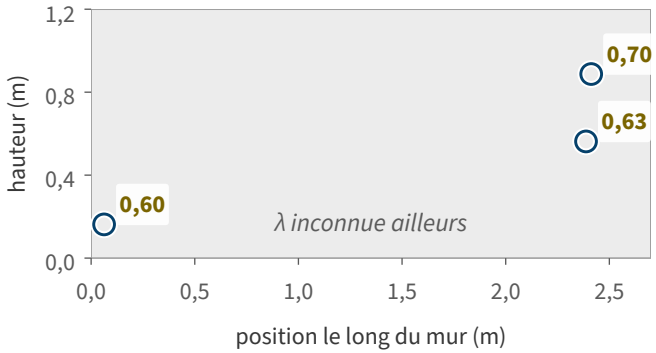
Mesure de la conductivité thermique λ sur trois éprouvettes carottées dans un mur en bauge.

Exemple pratique

Mesure de la conductivité thermique λ sur trois éprouvettes carottées dans un mur en bauge.

- Trois éprouvettes :

0,60 – 0,63 – 0,70 W/m·K.



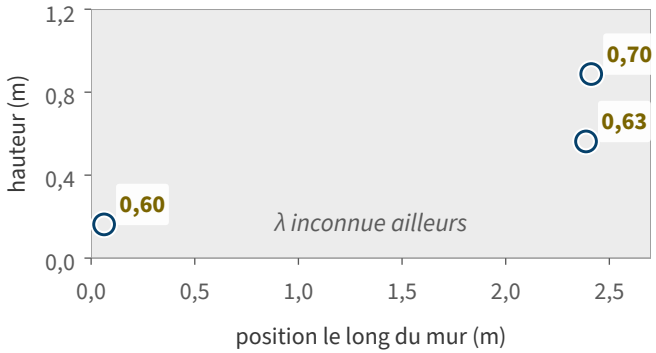
Exemple pratique

Mesure de la conductivité thermique λ sur trois éprouvettes carottées dans un mur en bauge.

- Trois éprouvettes :
0,60 – 0,63 – 0,70 W/m·K.

- Ce que l'on publie :
λ **= 0,64 ± 0,05 W/m·K** (*n* **= 3**)
  - moyenne $\hat{\mu}_\lambda = 0{,}64$
  - écart-type $\hat{\sigma}_\lambda = 0{,}05$

  le chapeau : valeur estimée



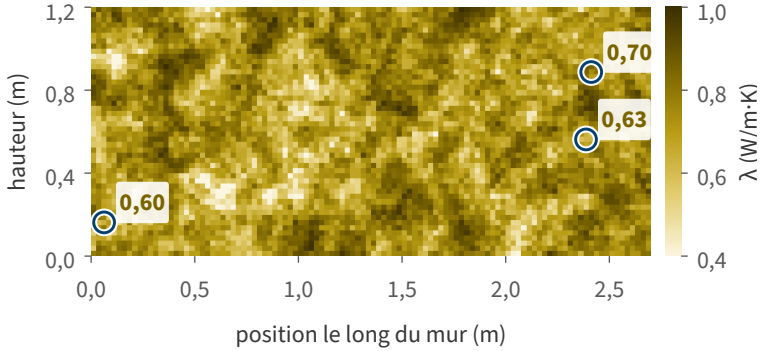
Exemple pratique

Mesure de la conductivité thermique λ sur trois éprouvettes carottées dans un mur en bauge.

Exemple jouet réaliste

- Trois éprouvettes :
0,60 – 0,63 – 0,70 W/m·K.
- Ce que l'on publie :
λ **= 0,64 ± 0,05 W/m·K** (n **= 3**)
  - moyenne $\hat{\mu}_\lambda$ = 0,64
  - écart-type $\hat{\sigma}_\lambda$ = 0,05

  le chapeau : valeur estimée
- **Imaginons que l'on connaisse λ en tout point du mur.** Que disent alors ces trois éprouvettes ?



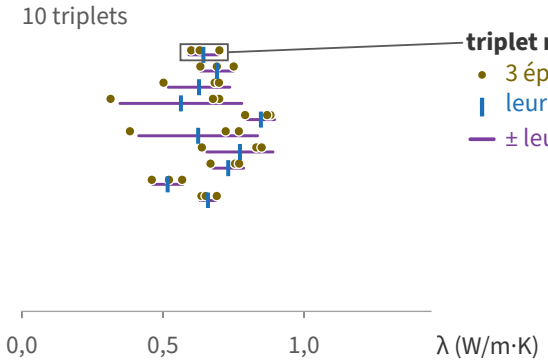
Exemple jouet : λ champ aléatoire de loi $f_\lambda = \mathcal{W}(k = 7{,}1\,;\ \eta = 0{,}76$ W/m·K), loi de Weibull de forme k et d'échelle η ; corrélation spatiale $\rho_\lambda(h) = \exp(-3h/0{,}22$ m) ; incertitude-type de mesure $u = 0{,}02$ W/m·K.

Exemple pratique : triplets du même mur

Trois autres éprouvettes du même mur auraient donné un autre résultat.

Exemple pratique : triplets du même mur

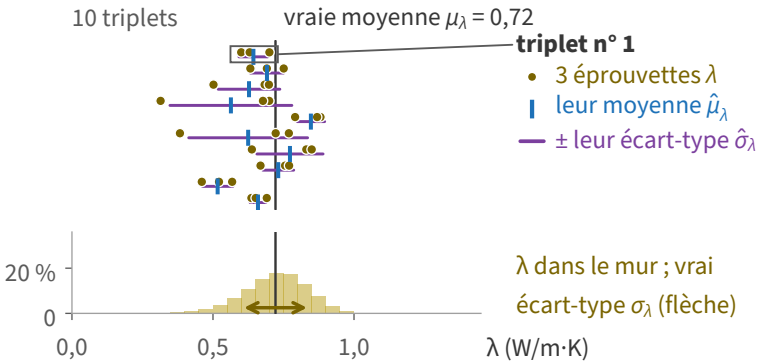
Trois autres éprouvettes du même mur auraient donné un autre résultat.



- 10 triplets, 10 résultats : ● $\hat{\mu}_\lambda$ et ● $\hat{\sigma}_\lambda$ sont des **estimateurs**.

Exemple pratique : triplets du même mur

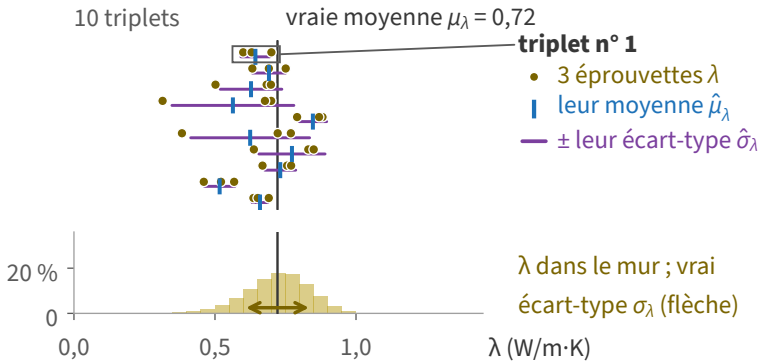
Trois autres éprouvettes du même mur auraient donné un autre résultat.



- 10 triplets, 10 résultats : ● $\hat{\mu}_\lambda$ et ● $\hat{\sigma}_\lambda$ sont des **estimateurs**.
- Le mur, lui, n'a pas changé : μ_λ et σ_λ sont les **paramètres vrais**.

Exemple pratique : triplets du même mur

Trois autres éprouvettes du même mur auraient donné un autre résultat.



- 10 triplets, 10 résultats : ● $\hat{\mu}_\lambda$ et ● $\hat{\sigma}_\lambda$ sont des **estimateurs**.
- Le mur, lui, n'a pas changé : μ_λ et σ_λ sont les **paramètres vrais**.

✔ Réponse

Trois éprouvettes, une moyenne, un écart-type : que sait-on ? **Moins qu'on ne croit : un autre triplet aurait donné un autre résultat.**

Objectif du cours : comprendre la variabilité, tester des hypothèses, planifier ses mesures

1. **Décrire la variabilité** — **Quelles notions pour décrire mes résultats ?**
2. **Planifier** — Combien d'éprouvettes ?
3. **Comparer** — Deux échantillons sont-ils différents ?
4. **Choisir une loi** — Quelle loi pour mes données ?
5. **Ajuster une tendance** — Que vaut mon ajustement ?
6. **Les réflexes** — Quels réflexes garder ?
7. **Ouverture** — *Et après ?* *(après le quiz, si le temps le permet)*

Incertitude de mesure et variabilité

Justesse et fidélité : deux qualités distinctes d'une mesure.

Incertitude de mesure et variabilité

Justesse et fidélité : deux qualités distinctes d'une mesure.

- **Justesse** : erreur systématique faible. Un biais se corrige par étalonnage.
- **Fidélité** : peu de dispersion entre mesures répétées.

juste et fidèle | **fidèle, pas juste** | **juste, pas fidèle** | **ni juste ni fidèle**



Schéma : six mesures répétées d'une même éprouvette par cible.

Incertitude de mesure et variabilité

Justesse et fidélité : deux qualités distinctes d'une mesure.

- **Justesse** : erreur systématique faible. Un biais se corrige par étalonnage.
- **Fidélité** : peu de dispersion entre mesures répétées.

juste et fidèle | **fidèle, pas juste** | **juste, pas fidèle** | **ni juste ni fidèle**



Schéma : six mesures répétées d'une même éprouvette par cible.

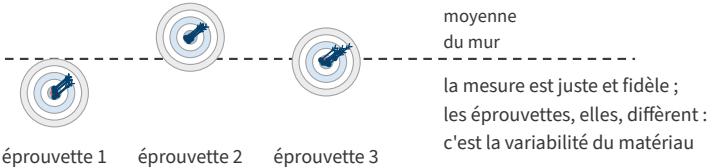
Une mesure : $\lambda = \lambda_{ép} + e$ ($\lambda_{ép}$: valeur de l'éprouvette, centre de la cible ; e : **erreur de mesure**).
Incertitude-type u : dispersion de e.

Incertitude de mesure et variabilité

Une mesure juste et fidèle ne supprime pas la variabilité du matériau : elle la caractérise.

Incertitude de mesure et variabilité

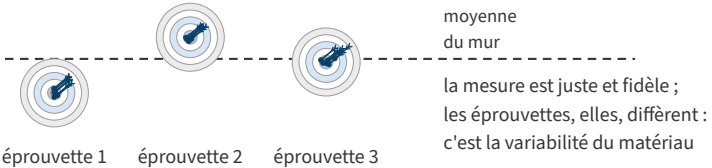
Une mesure juste et fidèle ne supprime pas la variabilité du matériau : elle la caractérise.



- Même juste et fidèle, la mesure trouve des éprouvettes différentes : $\lambda_{\text{ép}}$ varie, d'écart-type σ_{mat} (**variabilité du matériau**).

Incertitude de mesure et variabilité

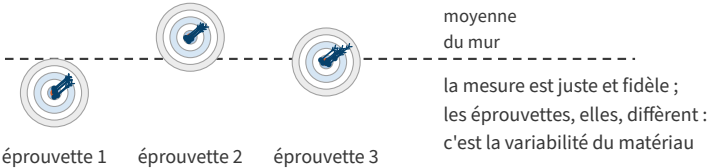
Une mesure juste et fidèle ne supprime pas la variabilité du matériau : elle la caractérise.



- Même juste et fidèle, la mesure trouve des éprouvettes différentes : $\lambda_{\text{ép}}$ varie, d'écart-type σ_{mat} (**variabilité du matériau**).
- Les variances s'additionnent : $\sigma_\lambda^2 = \sigma_{\text{mat}}^2 + u^2$. Ici $u^2 \approx 3\,\%$ de σ_λ^2 : **répéter la mesure réduit les 3 %, pas les 97 %** (les séparer : diapo 16).

Incertitude de mesure et variabilité

Une mesure juste et fidèle ne supprime pas la variabilité du matériau : elle la caractérise.



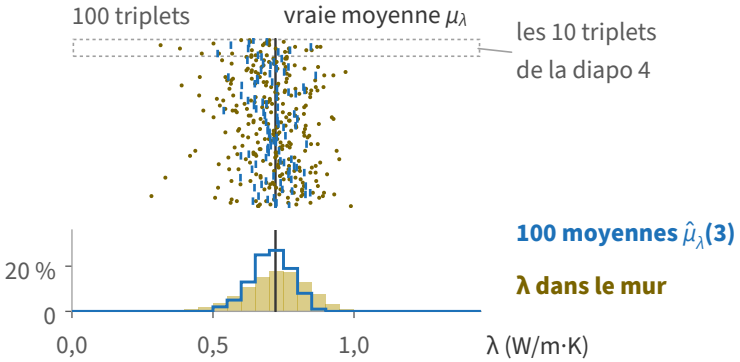
- Même juste et fidèle, la mesure trouve des éprouvettes différentes : $\lambda_{\text{ép}}$ varie, d'écart-type σ_{mat} (**variabilité du matériau**).
- Les variances s'additionnent : $\sigma_\lambda^2 = \sigma_{\text{mat}}^2 + u^2$. Ici $u^2 \approx 3\,\%$ de σ_λ^2 : **répéter la mesure réduit les 3 %, pas les 97 %** (les séparer : diapo 16).
- Incertitude de mesure : le **GUM**. Il ne dit ni comment caractériser σ_{mat}, ni combien d'éprouvettes prévoir, ni comment tester (diapo 36).

Écart-type ou erreur standard ?

L'écart-type décrit la variabilité des éprouvettes ; l'erreur standard, celle de la moyenne estimée.

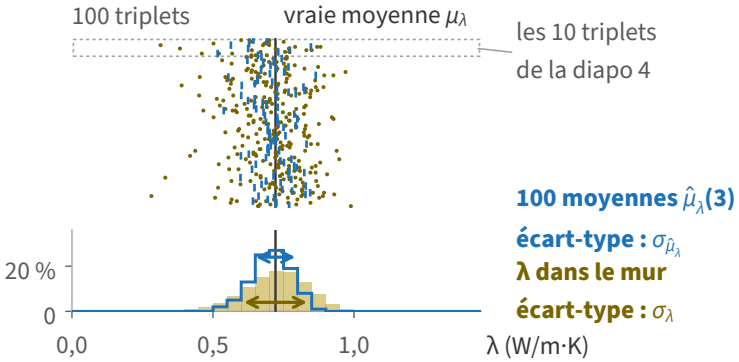
Écart-type ou erreur standard ?

L'écart-type décrit la variabilité des éprouvettes ; l'erreur standard, celle de la moyenne estimée.



Écart-type ou erreur standard ?

L'écart-type décrit la variabilité des éprouvettes ; l'erreur standard, celle de la moyenne estimée.

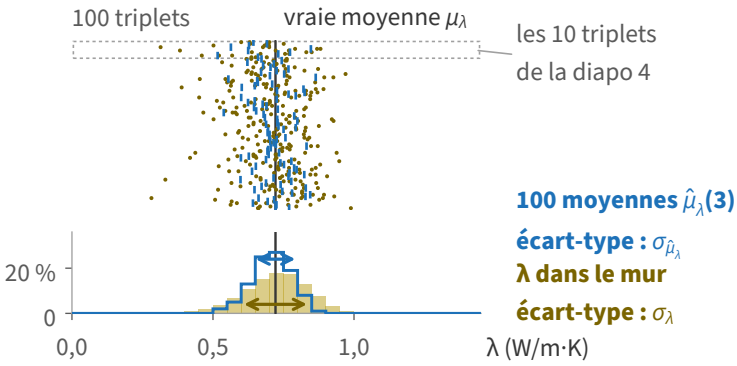


- **Écart-type** σ_λ : dispersion des éprouvettes ; ne diminue pas avec n.
- **Erreur standard** $\sigma_{\hat{\mu}_\lambda} = \sigma_\lambda/\sqrt{n}$: dispersion de $\hat{\mu}_\lambda$; diminue avec n. Estimée : $\hat{\sigma}_{\hat{\mu}_\lambda}$.

Tableur : `=ECARTYPE.STANDARD(plage) ; =ECARTYPE.STANDARD(plage)/RACINE(NB(plage))`
Python : `x.std(ddof=1), stats.sem(x)`

Écart-type ou erreur standard ?

L'écart-type décrit la variabilité des éprouvettes ; l'erreur standard, celle de la moyenne estimée.



- **Écart-type σ_λ** : dispersion des éprouvettes ; ne diminue pas avec n.
- **Erreur standard** $\sigma_{\hat{\mu}_\lambda} = \sigma_\lambda/\sqrt{n}$: dispersion de $\hat{\mu}_\lambda$; diminue avec n. Estimée : $\hat{\sigma}_{\hat{\mu}_\lambda}$.

Écart-type : les éprouvettes.
Erreur standard : la moyenne.

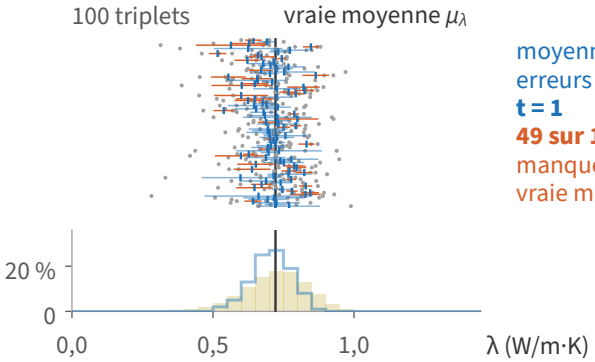
Tableur : `=ECARTYPE.STANDARD(plage) ; =ECARTYPE.STANDARD(plage)/RACINE(NB(plage))`
Python : `x.std(ddof=1), stats.sem(x)`

Intervalle de confiance d'un paramètre : exemple de la moyenne

L'IC à 95 % encadre la vraie moyenne 95 fois sur 100.

Intervalle de confiance d'un paramètre : exemple de la moyenne

L'IC à 95 % encadre la vraie moyenne 95 fois sur 100.



moyenne ± t erreurs standard
t = 1
49 sur 100
manquent la vraie moyenne

Quelles valeurs plausibles, avec quelle confiance ?

- ± 1 erreur standard $\hat{\sigma}_{\hat{\mu}_\lambda}$: près d'un intervalle sur deux manque μ_λ.

Intervalle de confiance d'un paramètre : exemple de la moyenne

L'IC à 95 % encadre la vraie moyenne 95 fois sur 100.

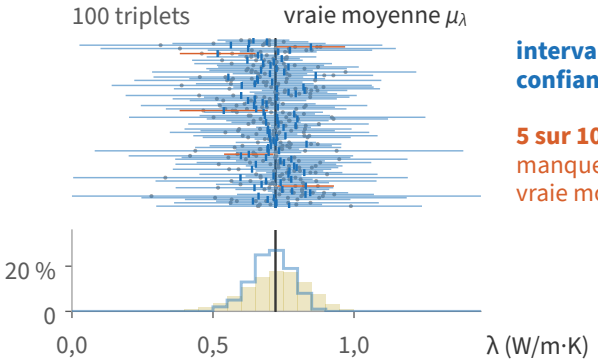
Quelles valeurs plausibles, avec quelle confiance ?

- ± 1 erreur standard $\hat{\sigma}_{\hat{\mu}_\lambda}$: près d'un intervalle sur deux manque μ_λ.
- ± t erreurs standard : plus t est grand, moins on manque μ_λ.

clic sur la figure : t suivant (Acrobat Reader) folioscope (annexe A4)

Intervalle de confiance d'un paramètre : exemple de la moyenne

L'IC à 95 % encadre la vraie moyenne 95 fois sur 100.



intervalle de confiance à 95 %

5 sur 100 manquent la vraie moyenne

Quelles valeurs plausibles, avec quelle confiance ?

- ± 1 erreur standard $\hat{\sigma}_{\hat{\mu}_\lambda}$: près d'un intervalle sur deux manque μ_λ.
- **Intervalle de confiance à 95 %** : $IC_{95}(\mu_\lambda) = \hat{\mu}_\lambda \pm t \times \hat{\sigma}_{\hat{\mu}_\lambda}$, t choisi pour 5 manqués sur 100.
- Moins d'éprouvettes : t plus grand, intervalle plus large.

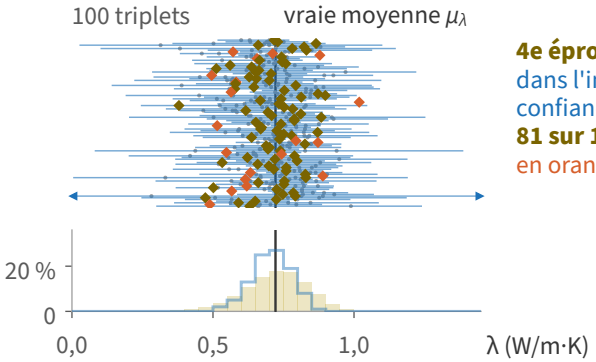
Tableur : `=INTERVALLE.CONFIANCE.STUDENT(0,05 ; ECARTYPE.STANDARD(plage) ; NB(plage))`
Python : `stats.t.ppf(0.975, len(x)-1) * stats.sem(x)` (demi-largeur, comme le tableur)

Intervalle de prédiction

L'IC à 95 % vise la vraie moyenne ; l'IP à 95 %, la prochaine éprouvette.

Intervalle de prédiction

L'IC à 95 % vise la vraie moyenne ; l'IP à 95 %, la prochaine éprouvette.

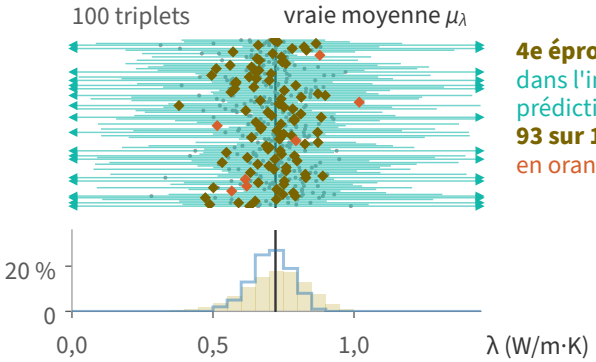


4e éprouvette
dans l'intervalle de confiance
81 sur 100
en orange : dehors

- ● 4e éprouvette dans $IC_{95}(\mu_\lambda)$: **81/100**.

Intervalle de prédiction

L'IC à 95 % vise la vraie moyenne ; l'IP à 95 %, la prochaine éprouvette.

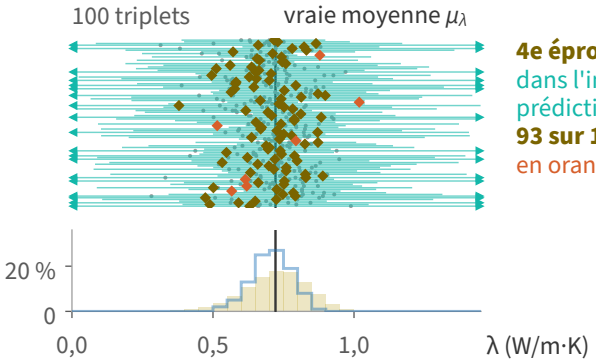


4e éprouvette
dans l'intervalle de prédiction
93 sur 100
en orange : dehors

- ● 4e éprouvette dans $IC_{95}(\mu_\lambda)$: **81/100**.
- ● 4e éprouvette dans $IP_{95}(\lambda)$: **93/100**.

Intervalle de prédiction

L'IC à 95 % vise la vraie moyenne ; l'IP à 95 %, la prochaine éprouvette.



4e éprouvette
dans l'intervalle de prédiction
93 sur 100
en orange : dehors

- ● 4e éprouvette dans $\text{IC}_{95}(\mu_\lambda)$: **81/100**.
- ● 4e éprouvette dans $\text{IP}_{95}(\lambda)$: **93/100**.
- **Intervalle de prédiction à 95 %** : il contient la prochaine éprouvette 95 fois sur 100 ; $\text{IP}_{95}(\lambda) = \hat{\mu}_\lambda \pm t \times \hat{\sigma}_\lambda \times \sqrt{1 + 1/n}$.

Tableur : `=LOI.STUDENT.INVERSE.BILATERALE(0,05 ; NB(plage)-1) * ECARTYPE.STANDARD(plage) * RACINE(1+1/NB(plage))`

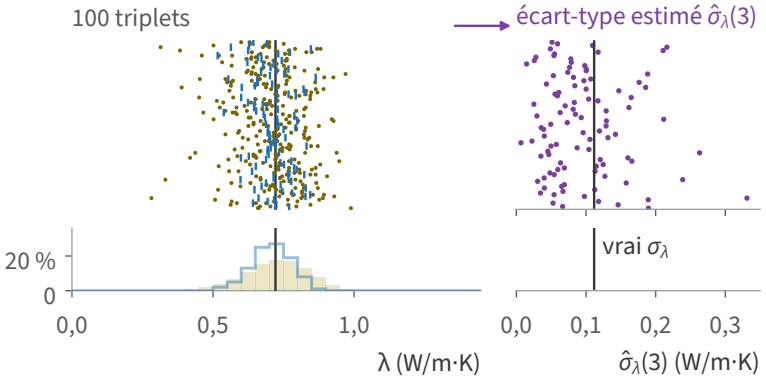
Python : `stats.t.ppf(0.975, len(x)-1) * x.std(ddof=1) * np.sqrt(1 + 1/len(x))`

Les poupées russes : l'écart-type lui-même est incertain

L'écart-type estimé $\hat{\sigma}_\lambda$ varie fortement d'un triplet à l'autre, et le CV avec lui.

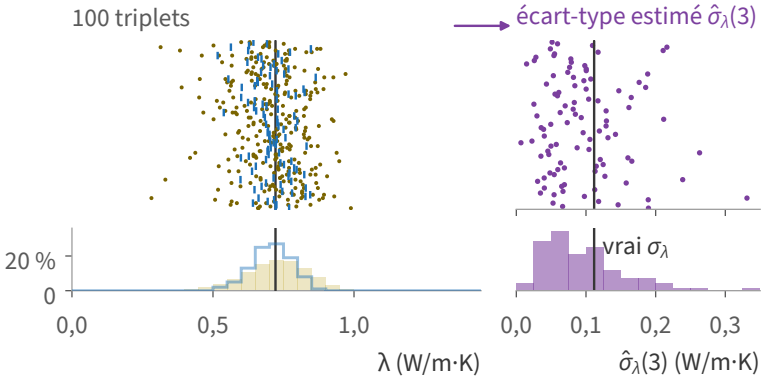
Les poupées russes : l'écart-type lui-même est incertain

L'écart-type estimé $\hat{\sigma}_\lambda$ varie fortement d'un triplet à l'autre, et le CV avec lui.



Les poupées russes : l'écart-type lui-même est incertain

L'écart-type estimé $\hat{\sigma}_\lambda$ varie fortement d'un triplet à l'autre, et le CV avec lui.

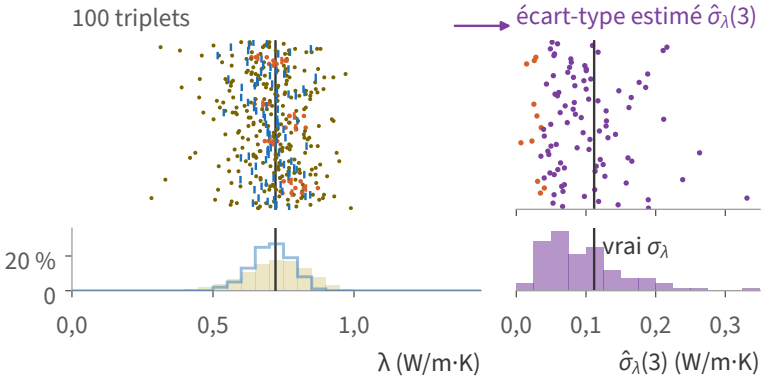


- Sur 100 triplets, ● $\hat{\sigma}_\lambda$ va de presque 0 à près de trois fois le vrai σ_λ.

À $n = 3$, l'écart-type lui-même est mal connu.

Les poupées russes : l'écart-type lui-même est incertain

L'écart-type estimé $\hat{\sigma}_\lambda$ varie fortement d'un triplet à l'autre, et le CV avec lui.



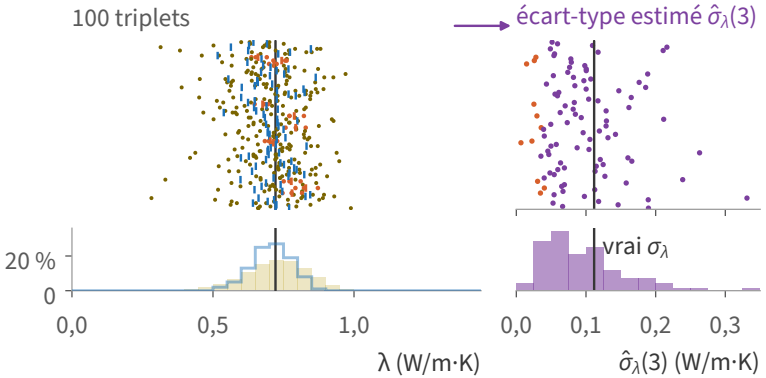
- Sur 100 triplets, ● $\hat{\sigma}_\lambda$ va de presque 0 à près de trois fois le vrai σ_λ.

À $n = 3$, l'écart-type lui-même est mal connu.

- **Coefficient de variation** CV = $\sigma_\lambda/\mu_\lambda$, ici 15,5 %. Son estimation en hérite : **11 triplets sur 100** affichent un CV sous 5 %.

Les poupées russes : l'écart-type lui-même est incertain

L'écart-type estimé $\hat{\sigma}_\lambda$ varie fortement d'un triplet à l'autre, et le CV avec lui.



- Sur 100 triplets, ● $\hat{\sigma}_\lambda$ va de presque 0 à près de trois fois le vrai σ_λ.

À $n = 3$, l'écart-type lui-même est mal connu.

- **Coefficient de variation** CV = $\sigma_\lambda/\mu_\lambda$, ici 15,5 %. Son estimation en hérite : **11 triplets sur 100** affichent un CV sous 5 %.

Cas publié : des écarts-types faibles, sur 3 éprouvettes par configuration, lus comme la preuve d'une bonne répétabilité et de mesures fiables.

Écrire ce qu'on a fait

Donner n, ce que vaut le « ± » et l'unité de répétition.

Écrire ce qu'on a fait

Donner n, ce que vaut le « ± » et l'unité de répétition.

Les mêmes trois éprouvettes, écrites de quatre façons :

0,64 ± 0,05 écart-type $\hat{\sigma}_\lambda$ ($n = 3$)

0,64 ± 0,03 erreur standard $\hat{\sigma}_{\hat{\mu}_\lambda}$ ($n = 3$)

0,64 ± 0,13 $\mathrm{IC}_{95}(\mu_\lambda)$ ($n = 3$)

0,64 ± 0,25 $\mathrm{IP}_{95}(\lambda)$ ($n = 3$)

Écrire ce qu'on a fait

Donner n, ce que vaut le « ± » et l'unité de répétition.

Les mêmes trois éprouvettes, écrites de quatre façons :

0,64 ± 0,05 écart-type $\hat{\sigma}_\lambda$ ($n = 3$)

0,64 ± 0,03 erreur standard $\hat{\sigma}_{\hat{\mu}_\lambda}$ ($n = 3$)

0,64 ± 0,13 $IC_{95}(\mu_\lambda)$ ($n = 3$)

0,64 ± 0,25 $IP_{95}(\lambda)$ ($n = 3$)

À écrire dans tout tableau

- ✔ n
- ✔ ce que vaut le « ± »
- ✔ l'**unité de répétition** (ce qu'on répète vraiment, l'éprouvette)

C'est aussi ce qu'on demande quand on relit un article.

Écrire ce qu'on a fait

Donner n, ce que vaut le « ± » et l'unité de répétition.

Les mêmes trois éprouvettes, écrites de quatre façons :

0,64 ± 0,05 écart-type $\hat{\sigma}_\lambda$ ($n = 3$)

0,64 ± 0,03 erreur standard $\hat{\sigma}_{\hat{\mu}_\lambda}$ ($n = 3$)

0,64 ± 0,13 $IC_{95}(\mu_\lambda)$ ($n = 3$)

0,64 ± 0,25 $IP_{95}(\lambda)$ ($n = 3$)

À écrire dans tout tableau

- ✔ n
- ✔ ce que vaut le « ± »
- ✔ l'**unité de répétition** (ce qu'on répète vraiment, l'éprouvette)

C'est aussi ce qu'on demande quand on relit un article.

✔ Réponse 1

Quelles notions pour décrire mes résultats ? IC_{95} **pour la vraie moyenne,** IP_{95} **pour la prochaine éprouvette, tous deux larges à** $n = 3$ **: écrire** n **et ce que vaut le « ± ».**

Objectif du cours : comprendre la variabilité, tester des hypothèses, planifier ses mesures

1. **Décrire la variabilité** Quelles notions pour décrire mes résultats ?

   ✔ **Réponse** IC_{95} pour la vraie moyenne, IP_{95} pour la prochaine éprouvette, tous deux larges à $n = 3$: écrire n et ce que vaut le « ± ».

2. **Planifier** **Combien d'éprouvettes ?**
3. **Comparer** Deux échantillons sont-ils différents ?
4. **Choisir une loi** Quelle loi pour mes données ?
5. **Ajuster une tendance** Que vaut mon ajustement ?
6. **Les réflexes** Quels réflexes garder ?
7. **Ouverture** *Et après ?* *(après le quiz, si le temps le permet)*

Pour estimer une moyenne et un écart-type

La moyenne se précise vite, l'écart-type lentement : *n* se décide avant l'essai.

Pour estimer une moyenne et un écart-type

La moyenne se précise vite, l'écart-type lentement : *n* se décide avant l'essai.

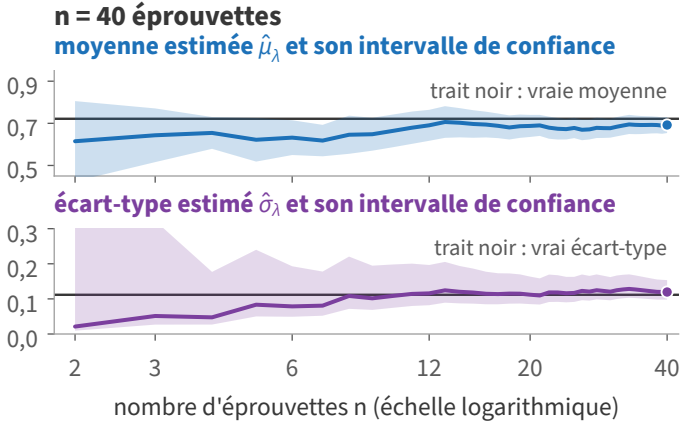
- Une campagne qui grandit sur le même mur.

clic sur la figure : *n* suivant (Acrobat Reader) folioscope (A1)

Pour estimer une moyenne et un écart-type

La moyenne se précise vite, l'écart-type lentement : n se décide avant l'essai.

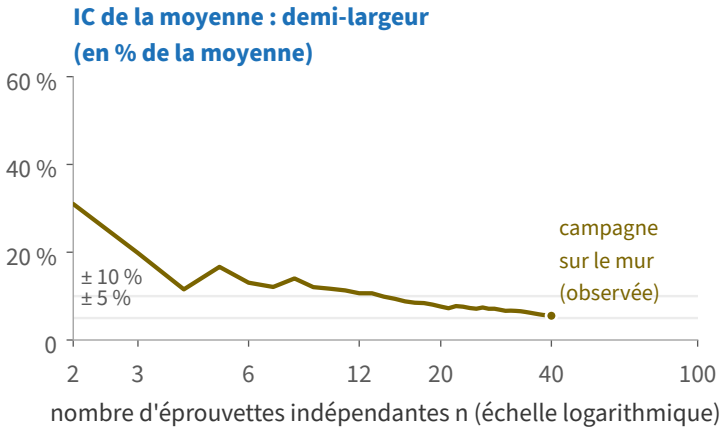
- Une campagne qui grandit sur le même mur.
- La moyenne se précise vite au début ; l'écart-type reste mal connu bien plus longtemps.



Pour estimer une moyenne et un écart-type

La moyenne se précise vite, l'écart-type lentement : *n* se décide avant l'essai.

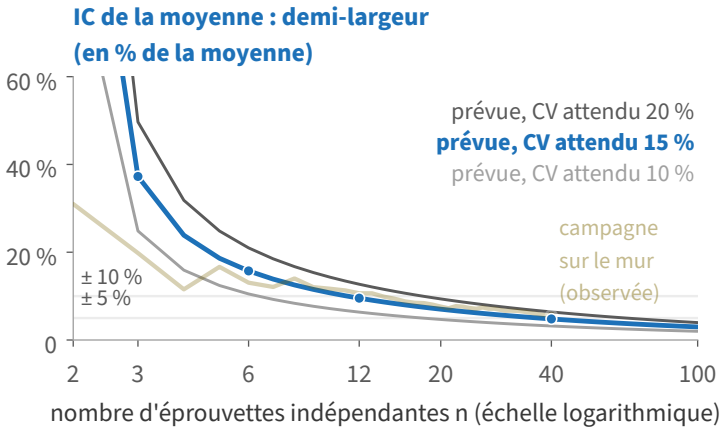
- Une campagne qui grandit sur le même mur.
- La moyenne se précise vite au début ; l'écart-type reste mal connu bien plus longtemps.



Pour estimer une moyenne et un écart-type

La moyenne se précise vite, l'écart-type lentement : *n* se décide avant l'essai.

- Une campagne qui grandit sur le même mur.
- La moyenne se précise vite au début ; l'écart-type reste mal connu bien plus longtemps.
- **Avant l'essai** : d'un CV attendu (littérature, essais préliminaires), l'abaque déduit le nombre d'éprouvettes pour la précision visée.

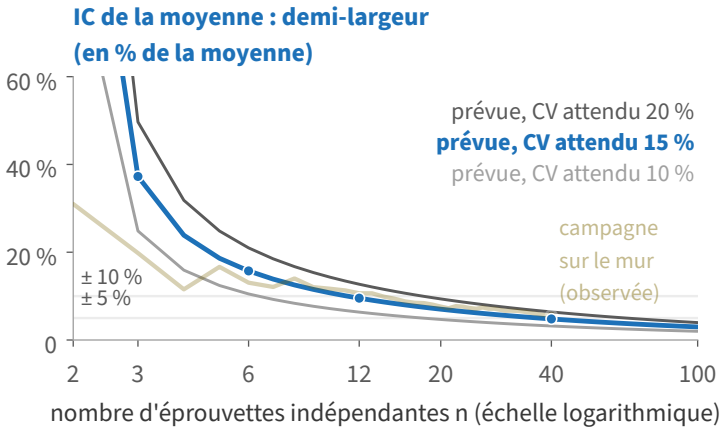


Pour estimer une moyenne et un écart-type

La moyenne se précise vite, l'écart-type lentement : *n* se décide avant l'essai.

- Une campagne qui grandit sur le même mur.
- La moyenne se précise vite au début ; l'écart-type reste mal connu bien plus longtemps.
- **Avant l'essai** : d'un CV attendu (littérature, essais préliminaires), l'abaque déduit le nombre d'éprouvettes pour la précision visée.

Cas publié : « au moins trois éprouvettes par condition, pour des résultats statistiquement représentatifs ».

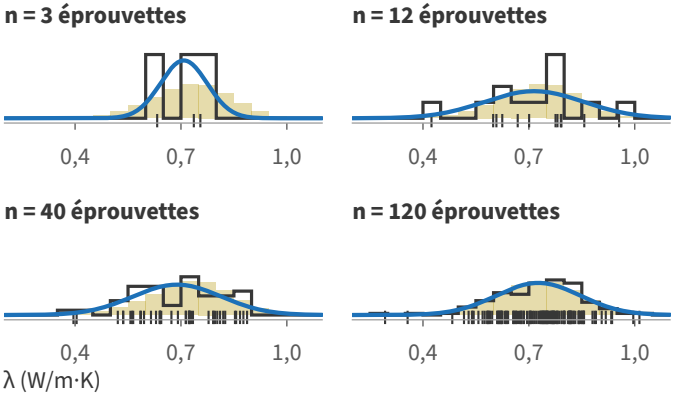


Pour caractériser la variabilité statistique

Estimation de la loi marginale et de ses paramètres : chacun demande plus d'éprouvettes que le précédent.

Pour caractériser la variabilité statistique

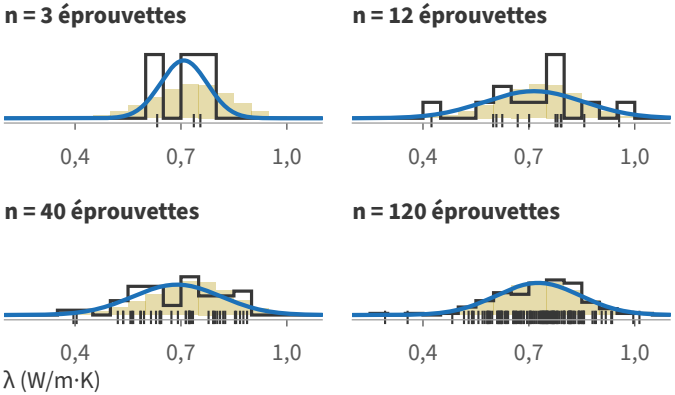
Estimation de la loi marginale et de ses paramètres : chacun demande plus d'éprouvettes que le précédent.



- Quatre campagnes sur le même mur, n = 3, 12, 40, 120.

Pour caractériser la variabilité statistique

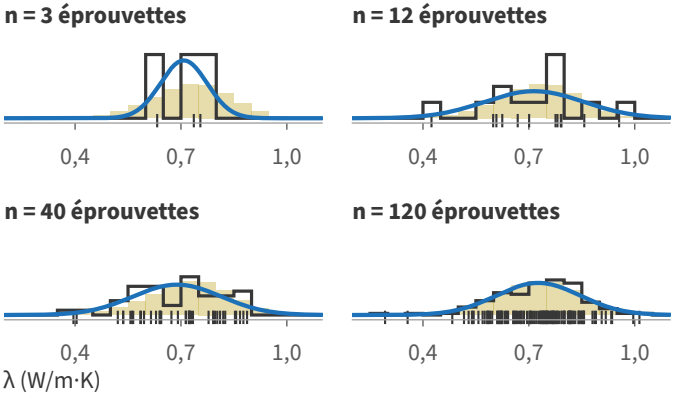
Estimation de la loi marginale et de ses paramètres : chacun demande plus d'éprouvettes que le précédent.



- Quatre campagnes sur le même mur, $n = 3, 12, 40, 120$.
- **La moyenne** μ_λ à ± 10 % : environ **12** éprouvettes (CV 15 %).
- **L'écart-type** σ_λ à ± 20 % : environ **70** éprouvettes.
- **La loi** : même avec 120 éprouvettes, on ne reconnaît la loi de Weibull du mur qu'une fois sur deux (cf. partie 4).

Pour caractériser la variabilité statistique

Estimation de la loi marginale et de ses paramètres : chacun demande plus d'éprouvettes que le précédent.



- Quatre campagnes sur le même mur, $n = 3, 12, 40, 120$.
- **La moyenne** μ_λ à ± 10 % : environ **12** éprouvettes (CV 15 %).
- **L'écart-type** σ_λ à ± 20 % : environ **70** éprouvettes.
- **La loi** : même avec 120 éprouvettes, on ne reconnaît la loi de Weibull du mur qu'une fois sur deux (cf. partie 4).

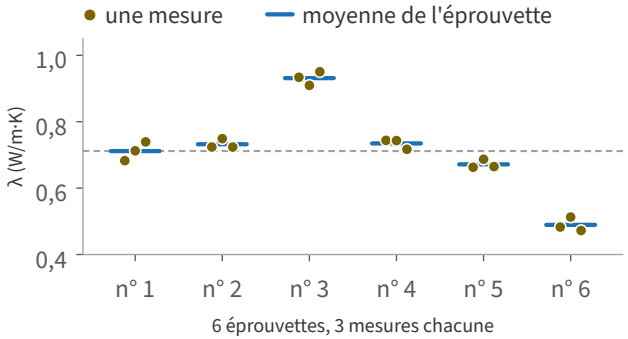
Trois éprouvettes donnent une moyenne grossière ; elles ne caractérisent pas la variabilité.

Séparer la mesure du matériau

Répéter la mesure de chaque éprouvette sépare u de σ_{mat}.

Séparer la mesure du matériau

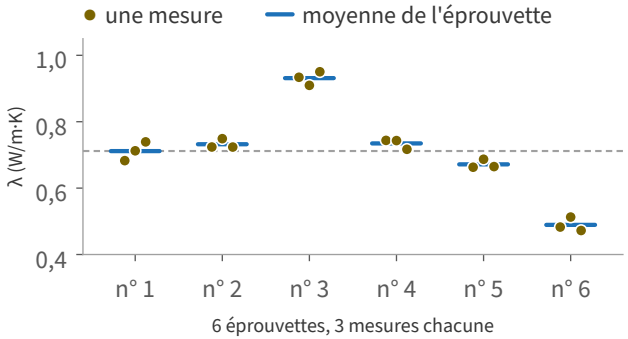
Répéter la mesure de chaque éprouvette sépare u de σ_{mat}.



- **Plan emboîté** : $n = 6$ éprouvettes, $r = 3$ mesures chacune ; chaque mesure refait tout le protocole (GUM, F.1.1.2).

Séparer la mesure du matériau

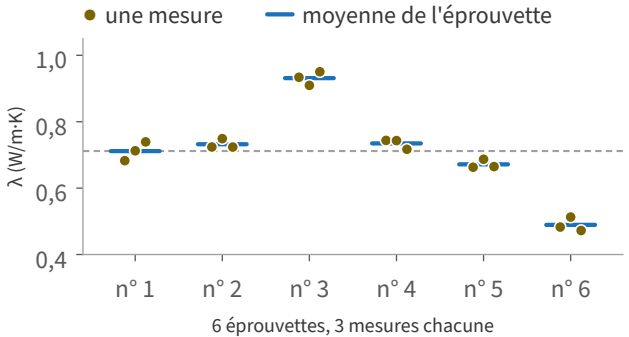
Répéter la mesure de chaque éprouvette sépare u de σ_{mat}.



- **Plan emboîté** : $n = 6$ éprouvettes, $r = 3$ mesures chacune ; chaque mesure refait tout le protocole (GUM, F.1.1.2).
- **Dans une éprouvette**, seule la mesure varie : $\hat{u} = 0{,}020$ W/m·K.

Séparer la mesure du matériau

Répéter la mesure de chaque éprouvette sépare u de σ_{mat}.



- **Plan emboîté** : $n = 6$ éprouvettes, $r = 3$ mesures chacune ; chaque mesure refait tout le protocole (GUM, F.1.1.2).
- **Dans une éprouvette**, seule la mesure varie : $\hat{u} = 0{,}020$ W/m·K.
- **Entre éprouvettes** (GUM, H.5) :

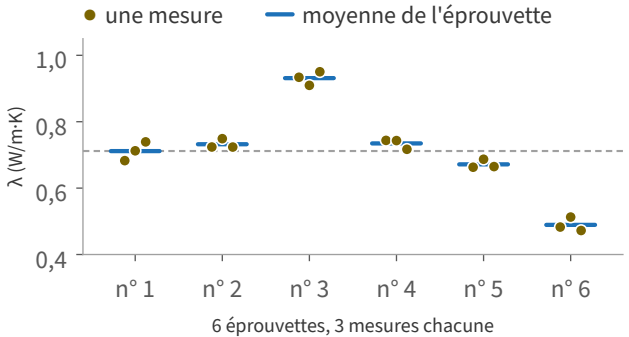
$$\hat{\sigma}^2_{mat} = \hat{\sigma}^2_{moy} - \hat{u}^2/r$$

$\hat{\sigma}^2_{moy}$: variance des n moyennes d'éprouvette → le matériau se lit sur n éprouvettes, pas sur $n \times r$ mesures.

Tableur : « Anova : un facteur » : $\hat{u}^2$ = CM intra ; $\hat{\sigma}^2_{mat}$ = (CM inter – $\hat{u}^2$)/r (CM : carré moyen)
Python : `u2 = x.var(1, ddof=1).mean(); s2_mat = x.mean(1).var(ddof=1) - u2/x.shape[1]`

Séparer la mesure du matériau

Répéter la mesure de chaque éprouvette sépare u de σ_{mat}.



- **Plan emboîté** : $n = 6$ éprouvettes, $r = 3$ mesures chacune ; chaque mesure refait tout le protocole (GUM, F.1.1.2).
- **Dans une éprouvette**, seule la mesure varie : $\hat{u} = 0{,}020$ W/m·K.
- **Entre éprouvettes** (GUM, H.5) :

$$\hat{\sigma}^2_{mat} = \hat{\sigma}^2_{moy} - \hat{u}^2/r$$

$\hat{\sigma}^2_{moy}$: variance des n moyennes d'éprouvette → le matériau se lit sur n éprouvettes, pas sur $n \times r$ mesures.

Les mesures répétées chiffrent u ; elles ne remplacent pas des éprouvettes (diapo suivante).

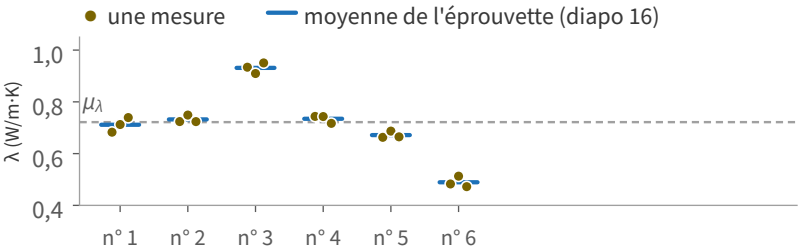
Tableur : « Anova : un facteur » : $\hat{u}^2$ = CM intra ; $\hat{\sigma}^2_{mat}$ = (CM inter – $\hat{u}^2$)/r (CM : carré moyen)
Python : `u2 = x.var(1, ddof=1).mean(); s2_mat = x.mean(1).var(ddof=1) - u2/x.shape[1]`

Attention à l'indépendance des éprouvettes

18 mesures sur 6 éprouvettes ne font pas 18 éprouvettes.

Attention à l'indépendance des éprouvettes

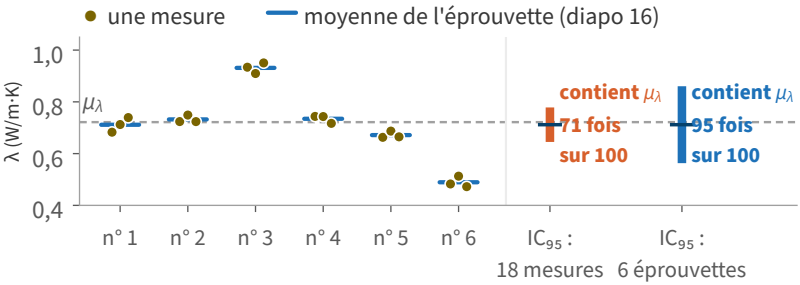
18 mesures sur 6 éprouvettes ne font pas 18 éprouvettes.



- Les données de la diapo 16.

Attention à l'indépendance des éprouvettes

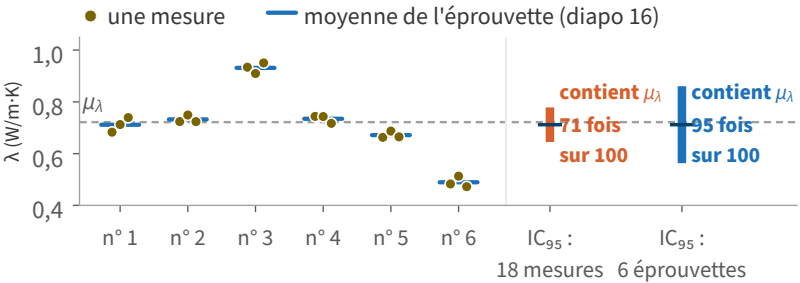
18 mesures sur 6 éprouvettes ne font pas 18 éprouvettes.



- Les données de la diapo 16.
- 18 mesures comptées « indépendantes » → IC deux fois trop étroit.

Attention à l'indépendance des éprouvettes

18 mesures sur 6 éprouvettes ne font pas 18 éprouvettes.

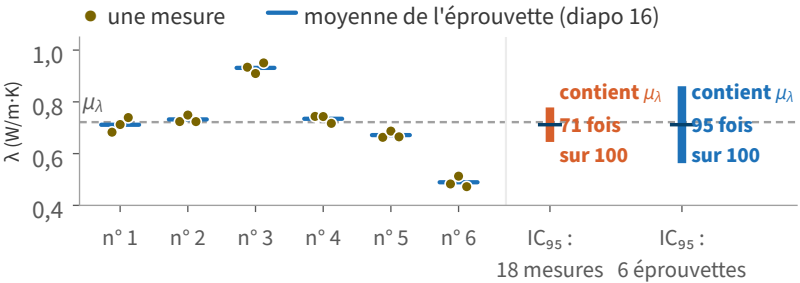


- Les données de la diapo 16.
- 18 mesures comptées « indépendantes » → IC deux fois trop étroit.

L'unité de répétition est ce qu'on répète vraiment : ici l'éprouvette, pas la mesure.

Attention à l'indépendance des éprouvettes

18 mesures sur 6 éprouvettes ne font pas 18 éprouvettes.



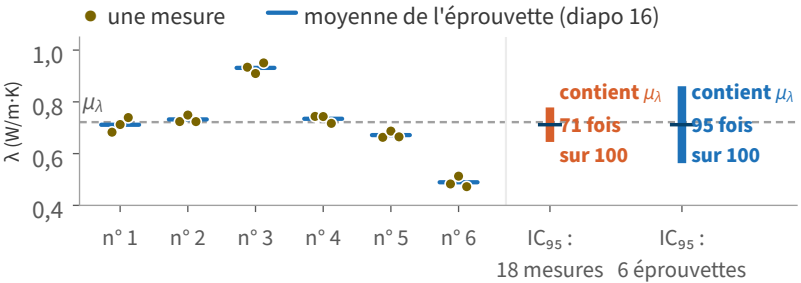
L'unité de répétition est ce qu'on répète vraiment : ici l'éprouvette, pas la mesure.

- Les données de la diapo 16.
- 18 mesures comptées « indépendantes » → IC deux fois trop étroit.

Cas publié : des éprouvettes voisines d'un même mur traitées comme indépendantes.

Attention à l'indépendance des éprouvettes

18 mesures sur 6 éprouvettes ne font pas 18 éprouvettes.



- Les données de la diapo 16.
- 18 mesures comptées « indépendantes » → IC deux fois trop étroit.

Cas publié : des éprouvettes voisines d'un même mur traitées comme indépendantes.

L'unité de répétition est ce qu'on répète vraiment : ici l'éprouvette, pas la mesure.

✔ Réponse 2

Combien d'éprouvettes ? **Avant l'essai, selon ce qu'on veut estimer et la précision visée ; seules comptent les éprouvettes indépendantes (unité de répétition).**

Objectif du cours : comprendre la variabilité, tester des hypothèses, planifier ses mesures

1. **Décrire la variabilité** Quelles notions pour décrire mes résultats ?
2. **Planifier** Combien d'éprouvettes ?

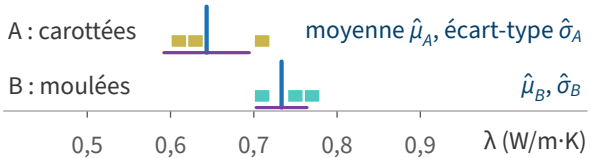
   ✔ **Réponse** Avant l'essai, selon ce qu'on veut estimer et la précision visée ; seules comptent les éprouvettes indépendantes (unité de répétition).
3. **Comparer** **Deux échantillons sont-ils différents ?**
4. **Choisir une loi** Quelle loi pour mes données ?
5. **Ajuster une tendance** Que vaut mon ajustement ?
6. **Les réflexes** Quels réflexes garder ?
7. **Ouverture** *Et après ?* *(après le quiz, si le temps le permet)*

Notion de test statistique

Comparer deux moyennes : l'écart observé est-il de ceux que produit le hasard ?

Notion de test statistique

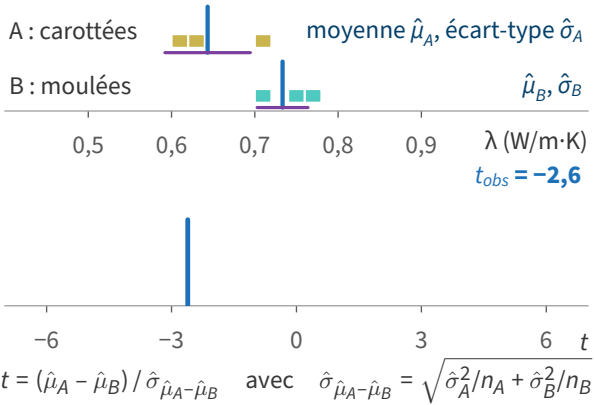
Comparer deux moyennes : l'écart observé est-il de ceux que produit le hasard ?



- **Hypothèse nulle** H0 : $\mu_A = \mu_B$.

Notion de test statistique

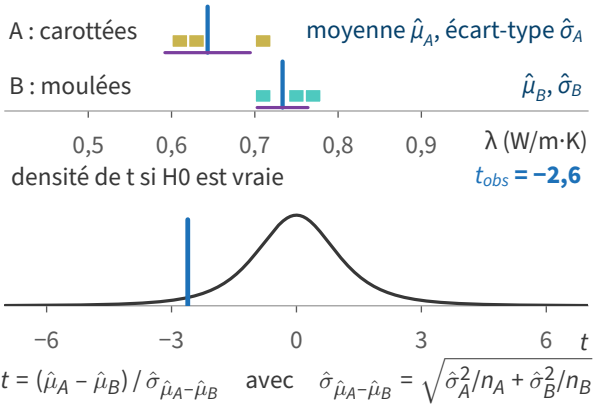
Comparer deux moyennes : l'écart observé est-il de ceux que produit le hasard ?



- **Hypothèse nulle** H0 : $\mu_A = \mu_B$.
- **Statistique de test** t : ici $t_{obs} = -2{,}6$.

Notion de test statistique

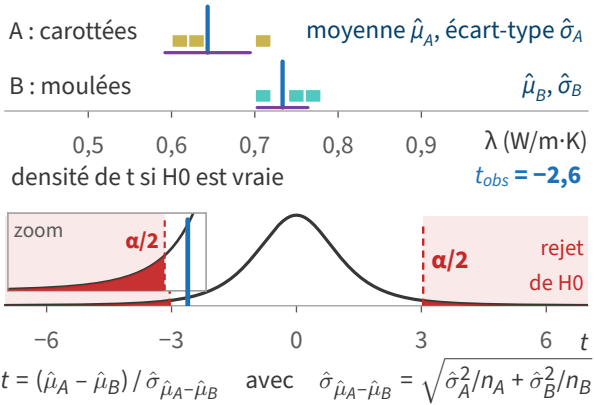
Comparer deux moyennes : l'écart observé est-il de ceux que produit le hasard ?



- **Hypothèse nulle** H0 : $\mu_A = \mu_B$.
- **Statistique de test** t : ici $t_{obs} = -2{,}6$.
- Si H0 est vraie : t suit une loi connue (la courbe).

Notion de test statistique

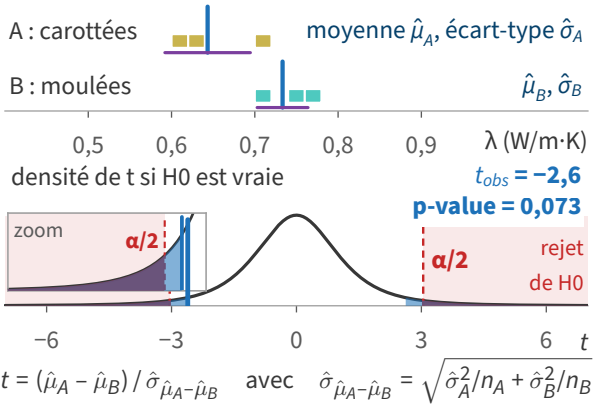
Comparer deux moyennes : l'écart observé est-il de ceux que produit le hasard ?



- **Hypothèse nulle** H0 : $\mu_A = \mu_B$.
- **Statistique de test** t : ici t_{obs} = −2,6.
- Si H0 est vraie : t suit une loi connue (la courbe).
- **Risque** α = P(rejeter H0 | H0 vraie) = 5 %, fixé d'avance.

Notion de test statistique

Comparer deux moyennes : l'écart observé est-il de ceux que produit le hasard ?

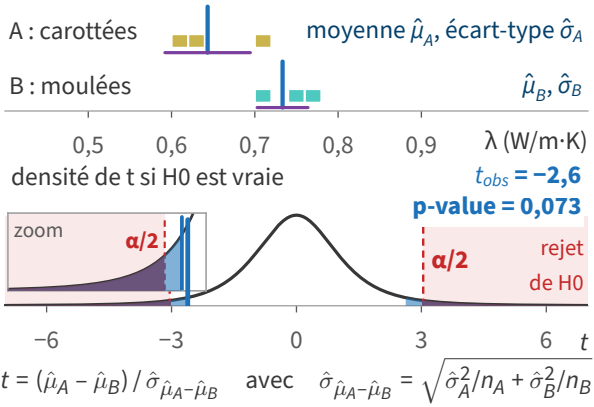


- **Hypothèse nulle** H0 : $\mu_A = \mu_B$.
- **Statistique de test** t : ici t_{obs} = −2,6.
- Si H0 est vraie : t suit une loi connue (la courbe).
- **Risque** α = P(rejeter H0 | H0 vraie) = 5 %, fixé d'avance.
- **p-value** = P($|T| \geq |t_{obs}|$ | H0) = **0,073** > α : H0 non rejetée… **ni prouvée**.

Tableur : `=T.TEST(plage1 ; plage2 ; 2 ; 3)` **Python** : `stats.ttest_ind(a, b, equal_var=False)`

Notion de test statistique

Comparer deux moyennes : l'écart observé est-il de ceux que produit le hasard ?



- **Hypothèse nulle** H0 : $\mu_A = \mu_B$.
- **Statistique de test** t : ici t_{obs} = −2,6.
- Si H0 est vraie : t suit une loi connue (la courbe).
- **Risque** α = P(rejeter H0 | H0 vraie) = 5 %, fixé d'avance.
- **p-value** = P($|T| \geq |t_{obs}|$ | H0) = **0,073** > α : H0 non rejetée… **ni prouvée**.

p-value : probabilité d'un écart au moins aussi net si seul le hasard jouait.

Tableur : `=T.TEST(plage1 ; plage2 ; 2 ; 3)` **Python** : `stats.ttest_ind(a, b, equal_var=False)`

Notion de test statistique

La p-value dépend de l'écart, mais aussi de la dispersion et de n.

Notion de test statistique

La p-value dépend de l'écart, mais aussi de la dispersion et de *n*.

rpsychologist.com/d3/nhst

folioscope (A3)

Ce que (ne) permet (pas) un test

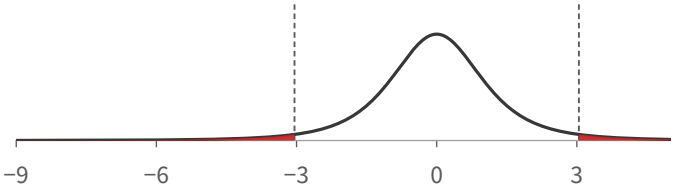
Avec peu d'éprouvettes, un test ne détecte que les grands effets.

Ce que (ne) permet (pas) un test

Avec peu d'éprouvettes, un test ne détecte que les grands effets.

Vraie différence : 0,04 W/m·K, soit 0,4 écart-type.

n = 3 éprouvettes par échantillon



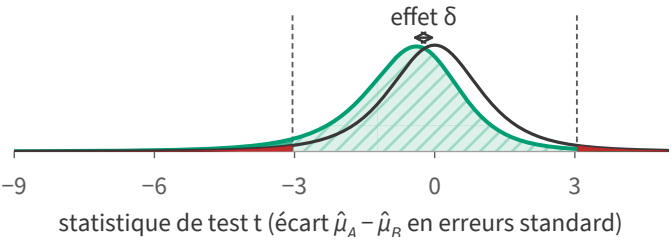
- H0 vraie (noir) : rejet à tort, probabilité α.

Ce que (ne) permet (pas) un test

Avec peu d'éprouvettes, un test ne détecte que les grands effets.

Vraie différence : 0,04 W/m·K, soit 0,4 écart-type. À $n = 3$, le test la détecte 6,5 fois sur 100.

n = 3 éprouvettes par échantillon **puissance 1 − β = 6,5 %**



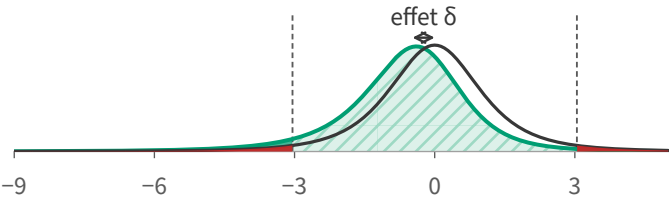
- H0 vraie (noir) : rejet à tort, probabilité α.
- Effet réel δ (vert) : manqué, probabilité β (vert pâle) ; **puissance** $1 - \beta$.

Ce que (ne) permet (pas) un test

Avec peu d'éprouvettes, un test ne détecte que les grands effets.

Vraie différence : 0,04 W/m·K, soit 0,4 écart-type. À $n = 3$, le test la détecte 6,5 fois sur 100.

n = 3 éprouvettes par échantillon **puissance 1 − β = 6,5 %**



	H0 vraie	H0 fausse (effet réel)
rejeter H0	faux positif : α	vrai positif : puissance $1 - \beta$
ne pas rejeter	vrai négatif : $1 - \alpha$	faux négatif : β

- H0 vraie (noir) : rejet à tort, probabilité α.
- Effet réel δ (vert) : manqué, probabilité β (vert pâle) ; **puissance** $1 - \beta$.

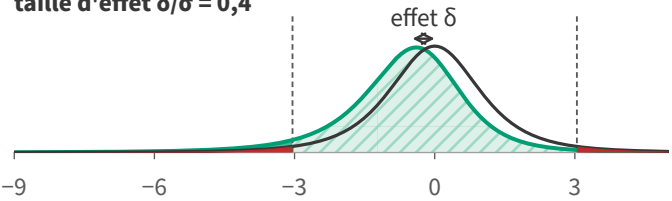
Ce que (ne) permet (pas) un test

Avec peu d'éprouvettes, un test ne détecte que les grands effets.

Vraie différence : 0,04 W/m·K, soit 0,4 écart-type. À $n = 3$, le test la détecte 6,5 fois sur 100.

n = 3 éprouvettes par échantillon **puissance 1 − β = 6,5 %**

taille d'effet δ/σ = 0,4



	H0 vraie	H0 fausse (effet réel)
rejeter H0	faux positif : α	vrai positif : puissance $1 - \beta$
ne pas rejeter	vrai négatif : $1 - \alpha$	faux négatif : β

- H0 vraie (noir) : rejet à tort, probabilité α.
- Effet réel δ (vert) : manqué, probabilité β (vert pâle) ; **puissance** $1 - \beta$.
- La puissance croît avec n et la **taille d'effet** δ/σ (ici l'écart vrai, connu car simulé).

Ce que (ne) permet (pas) un test

Avec peu d'éprouvettes, un test ne détecte que les grands effets.

Vraie différence : 0,04 W/m·K, soit 0,4 écart-type. À $n = 3$, le test la détecte 6,5 fois sur 100.

	H0 vraie	H0 fausse (effet réel)
rejeter H0	faux positif : α	vrai positif : puissance $1 - \beta$
ne pas rejeter	vrai négatif : $1 - \alpha$	faux négatif : β

- H0 vraie (noir) : rejet à tort, probabilité α.
- Effet réel δ (vert) : manqué, probabilité β (vert pâle) ; **puissance** $1 - \beta$.
- La puissance croît avec n et la **taille d'effet** δ/σ (ici l'écart vrai, connu car simulé).

clic sur la figure : image suivante

folioscope (A5)

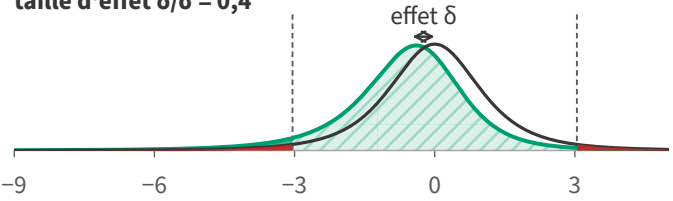
Ce que (ne) permet (pas) un test

Avec peu d'éprouvettes, un test ne détecte que les grands effets.

Vraie différence : 0,04 W/m·K, soit 0,4 écart-type. À $n = 3$, le test la détecte 6,5 fois sur 100.

n = 3 éprouvettes par échantillon **puissance 1 − β = 6,5 %**

taille d'effet δ/σ = 0,4



	H0 vraie	H0 fausse (effet réel)
rejeter H0	faux positif : α	vrai positif : puissance $1 - \beta$
ne pas rejeter	vrai négatif : $1 - \alpha$	faux négatif : β

- H0 vraie (noir) : rejet à tort, probabilité α.
- Effet réel δ (vert) : manqué, probabilité β (vert pâle) ; **puissance** $1 - \beta$.
- La puissance croît avec n et la **taille d'effet** δ/σ (ici l'écart vrai, connu car simulé).

« Non significatif » n'est pas « pas de différence ». La p-value ne mesure ni l'effet, ni la probabilité que H0 soit vraie.

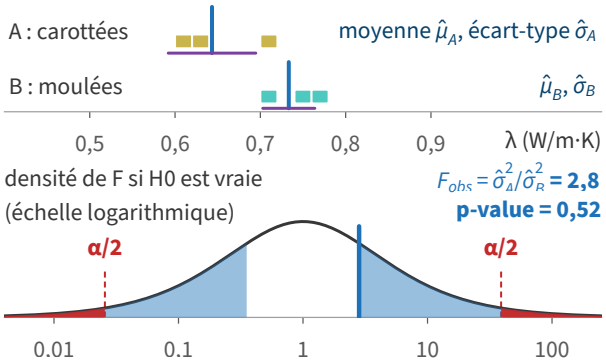
rpsychologist.com/d3/nhst

Comparer deux variances

À $n = 3$, le test F ne rejette H0 que si le rapport observé des variances dépasse 39 (ou 1/39).

Comparer deux variances

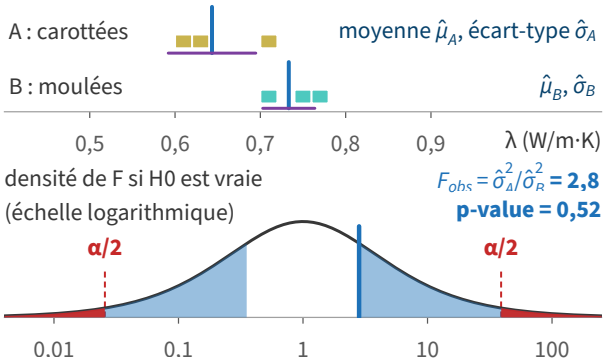
À $n = 3$, le test F ne rejette H0 que si le rapport observé des variances dépasse 39 (ou 1/39).



- H0 : $\sigma_A = \sigma_B$. Test F : $F = \hat{\sigma}_A^2/\hat{\sigma}_B^2 = 2{,}8$; $p = \mathbf{0{,}52}$: non rejetée.

Comparer deux variances

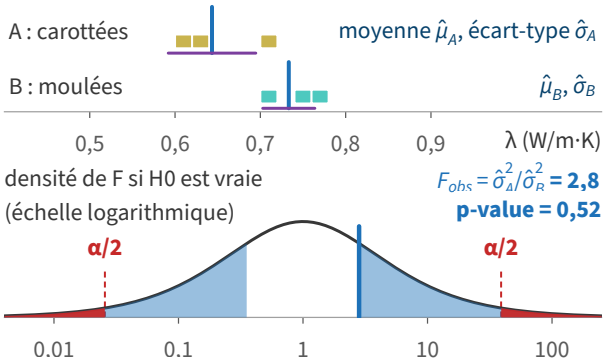
À $n = 3$, le test F ne rejette H0 que si le rapport observé des variances dépasse 39 (ou 1/39).



- H0 : $\sigma_A = \sigma_B$. Test F : $F = \hat{\sigma}_A^2/\hat{\sigma}_B^2 = 2{,}8$; p = **0,52** : non rejetée.
- **Taille d'effet** (rapport des variances) : observée 2,8 ; vraie 1,8.

Comparer deux variances

À $n = 3$, le test F ne rejette H0 que si le rapport observé des variances dépasse 39 (ou 1/39).



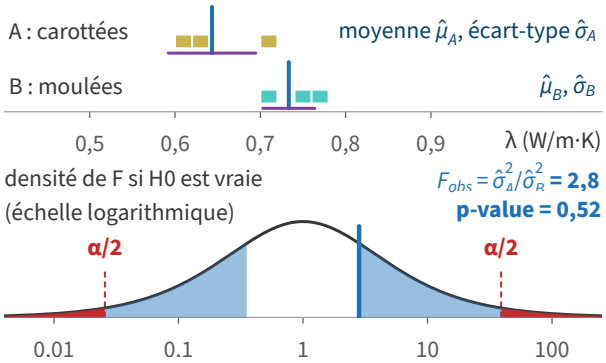
- H0 : $\sigma_A = \sigma_B$. Test F : $F = \hat{\sigma}_A^2/\hat{\sigma}_B^2 = 2{,}8$; p = **0,52** : non rejetée.
- **Taille d'effet** (rapport des variances) : observée 2,8 ; vraie 1,8.
- **Puissance** à $n = 3$ (rapport vrai) : 5,8 %.

Tableur : `=F.TEST(plage1 ; plage2)`

Python :
```
F = a.var(ddof=1) / b.var(ddof=1); d = len(a)-1, len(b)-1
2*min(stats.f.sf(F, *d), stats.f.cdf(F, *d))
```

Comparer deux variances

À $n = 3$, le test F ne rejette H0 que si le rapport observé des variances dépasse 39 (ou 1/39).



- H0 : $\sigma_A = \sigma_B$. Test F : $F = \hat{\sigma}_A^2/\hat{\sigma}_B^2 = 2{,}8$; p = **0,52** : non rejetée.
- **Taille d'effet** (rapport des variances) : observée 2,8 ; vraie 1,8.
- **Puissance** à $n = 3$ (rapport vrai) : 5,8 %.

À $n = 3$, ni moyennes ni variances ne diffèrent significativement : ces tests manquent de puissance.

Tableur : `=F.TEST(plage1 ; plage2)`

Python :
```
F = a.var(ddof=1) / b.var(ddof=1); d = len(a)-1, len(b)-1
2*min(stats.f.sf(F, *d), stats.f.cdf(F, *d))
```

Quel test pour quelle question ?

Le test se choisit d'après la question, avant l'essai.

Quel test pour quelle question ?

Le test se choisit d'après la question, avant l'essai.

Question	Test (hypothèses)	Tableur	Python (scipy.stats)
Deux moyennes	Welch ; Student si variances égales	`T.TEST(a;b;2;3)`	`ttest_ind(a, b, equal_var=False)`
Position (médianes), sans hypothèse de loi	Mann-Whitney (rangs)	—	`mannwhitneyu(a, b)`
Plusieurs moyennes	analyse de variance (ANOVA)	utilitaire d'analyse	`f_oneway(a, b, c)`
Deux variances	F (lois normales), Levene	`F.TEST(a;b)`	`levene(a, b)`

Quel test pour quelle question ?

Le test se choisit d'après la question, avant l'essai.

Question	Test (hypothèses)	Tableur	Python (scipy.stats)
Deux moyennes	Welch ; Student si variances égales	`T.TEST(a;b;2;3)`	`ttest_ind(a, b, equal_var=False)`
Position (médianes), sans hypothèse de loi	Mann-Whitney (rangs)	—	`mannwhitneyu(a, b)`
Plusieurs moyennes	analyse de variance (ANOVA)	utilitaire d'analyse	`f_oneway(a, b, c)`
Deux variances	F (lois normales), Levene	`F.TEST(a;b)`	`levene(a, b)`

Bonnes pratiques (d'après Mascha et Vetter, *Anesth. Analg.*, 2018)

- **Avant l'essai** : α (souvent 0,05), puissance visée (80 %, mieux 90 %), plus petit effet qui compte (en unités physiques) → n.
- **Dans l'article** : test, n par échantillon, p-value, effet observé et son IC ; pas de puissance calculée sur l'effet observé.

Quel test pour quelle question ?

Le test se choisit d'après la question, avant l'essai.

Question	Test (hypothèses)	Tableur	Python (scipy.stats)
Deux moyennes	Welch ; Student si variances égales	`T.TEST(a;b;2;3)`	`ttest_ind(a, b, equal_var=False)`
Position (médianes), sans hypothèse de loi	Mann-Whitney (rangs)	—	`mannwhitneyu(a, b)`
Plusieurs moyennes	analyse de variance (ANOVA)	utilitaire d'analyse	`f_oneway(a, b, c)`
Deux variances	F (lois normales), Levene	`F.TEST(a;b)`	`levene(a, b)`

Bonnes pratiques (d'après Mascha et Vetter, *Anesth. Analg.*, 2018)

- **Avant l'essai** : α (souvent 0,05), puissance visée (80 %, mieux 90 %), plus petit effet qui compte (en unités physiques) → n.
- **Dans l'article** : test, n par échantillon, p-value, effet observé et son IC ; pas de puissance calculée sur l'effet observé.

✔ Réponse 3

Deux échantillons sont-ils différents ? **À $n = 3$, un test ne voit que les écarts énormes : « non significatif » n'est pas « pas de différence ».**

Objectif du cours : comprendre la variabilité, tester des hypothèses, planifier ses mesures

1. **Décrire la variabilité** Quelles notions pour décrire mes résultats ?
2. **Planifier** Combien d'éprouvettes ?
3. **Comparer** Deux échantillons sont-ils différents ?

   ✔ **Réponse** À $n = 3$, un test ne voit que les écarts énormes : « non significatif » n'est pas « pas de différence ».

4. **Choisir une loi** **Quelle loi pour mes données ?**
5. **Ajuster une tendance** Que vaut mon ajustement ?
6. **Les réflexes** Quels réflexes garder ?
7. **Ouverture** *Et après ?* *(après le quiz, si le temps le permet)*

Ajuster une loi : le maximum de vraisemblance

18 éprouvettes du mur : on retient les paramètres qui les rendent les plus vraisemblables.

Ajuster une loi : le maximum de vraisemblance

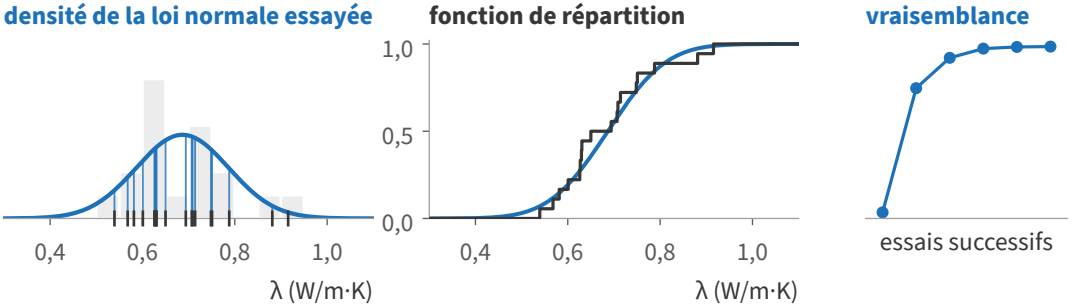
18 éprouvettes du mur : on retient les paramètres qui les rendent les plus vraisemblables.

- **Vraisemblance** : le produit des hauteurs des traits bleus. On retient les paramètres qui la rendent maximale : c'est le **maximum de vraisemblance**.

clic sur la figure : essai suivant (Acrobat Reader) folioscope (A6)

Ajuster une loi : le maximum de vraisemblance

18 éprouvettes du mur : on retient les paramètres qui les rendent les plus vraisemblables.



- **Vraisemblance** : le produit des hauteurs des traits bleus. On retient les paramètres qui la rendent maximale : c'est le **maximum de vraisemblance**.
- Loi normale : on retrouve la moyenne ; l'écart-type est calculé avec n, et non $n - 1$.

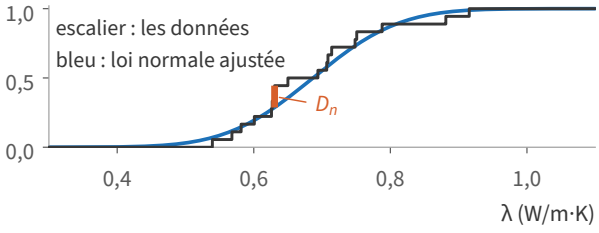
Tableur : `Solveur : maximiser la somme des LN(LOI.NORMALE.N(x ; mu ; sigma ; FAUX))`
Python : `stats.norm.fit(x) ; stats.weibull_min.fit(x, floc=0)`

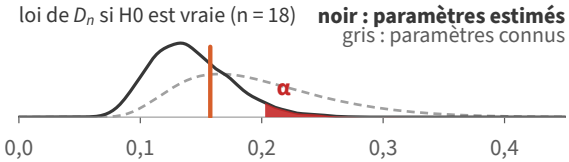
Tester un ajustement

Un test d'ajustement ne rejette une loi fausse qu'avec beaucoup d'éprouvettes.

Tester un ajustement

Un test d'ajustement ne rejette une loi fausse qu'avec beaucoup d'éprouvettes.

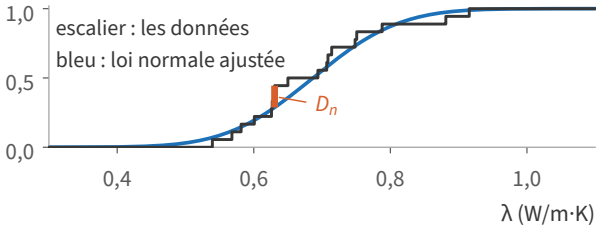


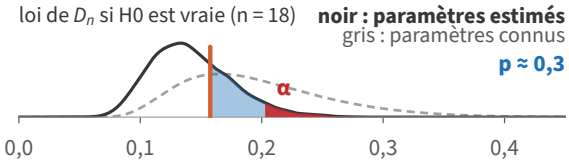


- Test de **Kolmogorov-Smirnov** : D_n = plus grand écart vertical entre l'escalier des données et la loi.
- H0 : « les données suivent cette loi ».

Tester un ajustement

Un test d'ajustement ne rejette une loi fausse qu'avec beaucoup d'éprouvettes.





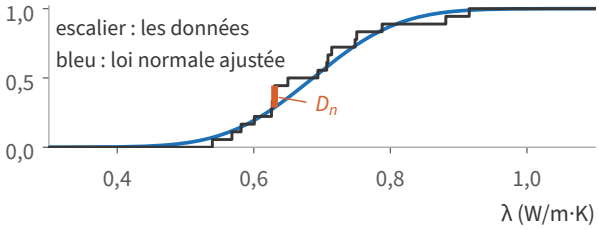
- Test de **Kolmogorov-Smirnov** : D_n = plus grand écart vertical entre l'escalier des données et la loi.
- H0 : « les données suivent cette loi ».
  $p \approx$ **0,3** : non rejetée (paramètres estimés : version de Lilliefors).

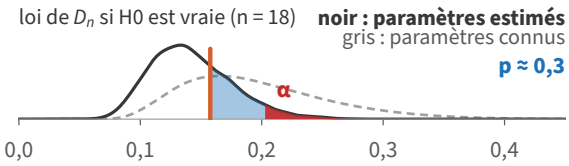
Tableur : –

Python : `from statsmodels.stats.diagnostic import lilliefors ; lilliefors(x)`

Tester un ajustement

Un test d'ajustement ne rejette une loi fausse qu'avec beaucoup d'éprouvettes.





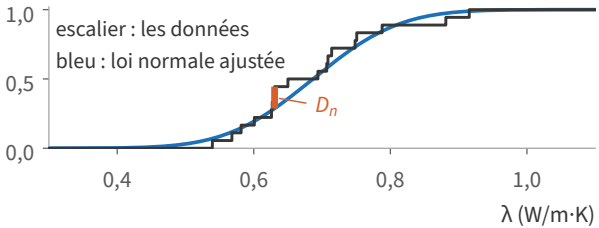
- Test de **Kolmogorov-Smirnov** : D_n = plus grand écart vertical entre l'escalier des données et la loi.
- H0 : « les données suivent cette loi ». $p \approx$ **0,3** : non rejetée (paramètres estimés : version de Lilliefors).
- **Puissance** contre la loi du mur : < 10 % (n = 18) ; ≈ 20 % (n = 132).

Tableur : –

Python : `from statsmodels.stats.diagnostic import lilliefors ; lilliefors(x)`

Tester un ajustement

Un test d'ajustement ne rejette une loi fausse qu'avec beaucoup d'éprouvettes.



- Test de **Kolmogorov-Smirnov** : D_n = plus grand écart vertical entre l'escalier des données et la loi.
- H0 : « les données suivent cette loi ». $p \approx$ **0,3** : non rejetée (paramètres estimés : version de Lilliefors).
- **Puissance** contre la loi du mur : < 10 % (n = 18) ; ≈ 20 % (n = 132).

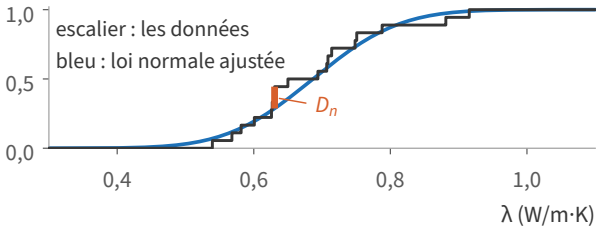
Ne pas rejeter une loi ne la prouve pas.

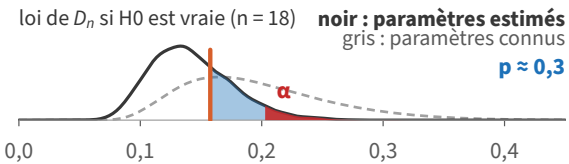
Tableur : –

Python : `from statsmodels.stats.diagnostic import lilliefors ; lilliefors(x)`

Tester un ajustement

Un test d'ajustement ne rejette une loi fausse qu'avec beaucoup d'éprouvettes.





- Test de **Kolmogorov-Smirnov** : D_n = plus grand écart vertical entre l'escalier des données et la loi.
- H0 : « les données suivent cette loi ». $p \approx$ **0,3** : non rejetée (paramètres estimés : version de Lilliefors).
- **Puissance** contre la loi du mur : < 10 % (n = 18) ; ≈ 20 % (n = 132).

Ne pas rejeter une loi ne la prouve pas.

Cas publié : loi normale « préférée » sur la foi de ce seul test.

Tableur : –

Python : `from statsmodels.stats.diagnostic import lilliefors ; lilliefors(x)`

Deux lois, aucune rejetée : les comparer

Reconnaître une loi demande beaucoup d'éprouvettes.

Deux lois, aucune rejetée : les comparer

Reconnaître une loi demande beaucoup d'éprouvettes.

18 éprouvettes : la normale un peu préférée…

… alors que le mur suit une loi de Weibull



≤ 2 : comparable	4 à 7 : nettement moins	> 10 : écartée

18 éprouvettes : écart 2,8 ▲

échelle : écart d'AICc entre les deux lois (repères de Burnham et Anderson, 2004)

AICc (critère d'Akaike corrigé), le plus petit gagne ; k paramètres, $\hat{L}$ vraisemblance maximale :

$$\mathrm{AICc} = 2k - 2\ln\hat{L} + 2k(k+1)/(n-k-1)$$

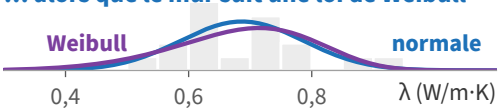
l'AICc selon n (annexe A7)

Deux lois, aucune rejetée : les comparer

Reconnaître une loi demande beaucoup d'éprouvettes.

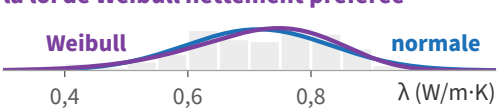
18 éprouvettes : la normale un peu préférée…

… alors que le mur suit une loi de Weibull



132 points (une campagne complète du mur) :

la loi de Weibull nettement préférée



≤ 2 : comparable | **4 à 7 : nettement moins** | **> 10 : écartée**

18 éprouvettes : écart 2,8 ▲ **▲ 132 points : écart 5,0**

échelle : écart d'AICc entre les deux lois (repères de Burnham et Anderson, 2004)

AICc (critère d'Akaike corrigé), le plus petit gagne ; k paramètres, $\hat{L}$ vraisemblance maximale :

$$\mathrm{AICc} = 2k - 2\ln\hat{L} + 2k(k+1)/(n-k-1)$$

18 éprouvettes : pas de décision 3 fois sur 4, mauvaise loi 1 fois sur 10.

Écrire « non rejetée », pas « validée ».

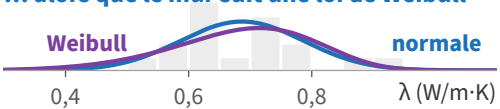
l'AICc selon n (annexe A7)

Deux lois, aucune rejetée : les comparer

Reconnaître une loi demande beaucoup d'éprouvettes.

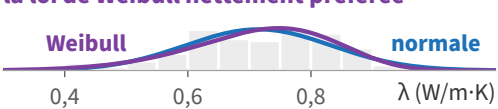
18 éprouvettes : la normale un peu préférée…

… alors que le mur suit une loi de Weibull



132 points (une campagne complète du mur) :

la loi de Weibull nettement préférée



≤ 2 : comparable | **4 à 7 : nettement moins** | **> 10 : écartée**

18 éprouvettes : écart 2,8 ▲ **▲ 132 points : écart 5,0**

échelle : écart d'AICc entre les deux lois (repères de Burnham et Anderson, 2004)

AICc (critère d'Akaike corrigé), le plus petit gagne ; k paramètres, $\hat{L}$ vraisemblance maximale :

$$\mathrm{AICc} = 2k - 2\ln\hat{L} + 2k(k+1)/(n-k-1)$$

18 éprouvettes : pas de décision 3 fois sur 4, mauvaise loi 1 fois sur 10.

Écrire « non rejetée », pas « validée ».

✔ Réponse 4

Quelle loi pour mes données ? **Souvent, plusieurs lois conviennent : ne pas rejeter une loi n'est pas la prouver ; on compare des modèles (AICc).**

Objectif du cours : comprendre la variabilité, tester des hypothèses, planifier ses mesures

1. **Décrire la variabilité** Quelles notions pour décrire mes résultats ?
2. **Planifier** Combien d'éprouvettes ?
3. **Comparer** Deux échantillons sont-ils différents ?
4. **Choisir une loi** Quelle loi pour mes données ?

   ✔ **Réponse** Souvent, plusieurs lois conviennent : ne pas rejeter une loi n'est pas la prouver ; on compare des modèles (AICc).

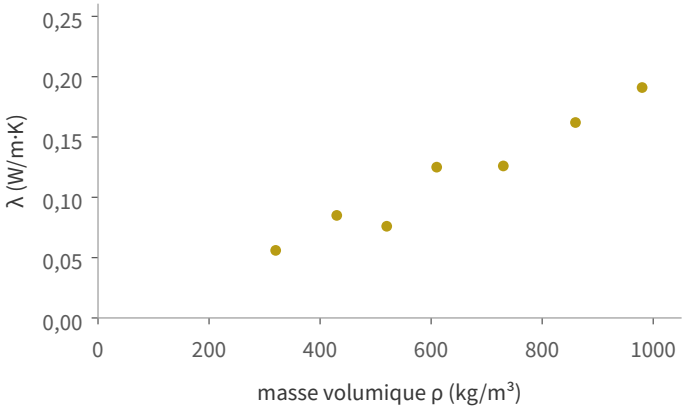
5. **Ajuster une tendance** **Que vaut mon ajustement ?**
6. **Les réflexes** Quels réflexes garder ?
7. **Ouverture** *Et après ?* *(après le quiz, si le temps le permet)*

L’incertitude d’une droite

Sept formulations de terre allégée : la droite ajustée est elle-même incertaine.

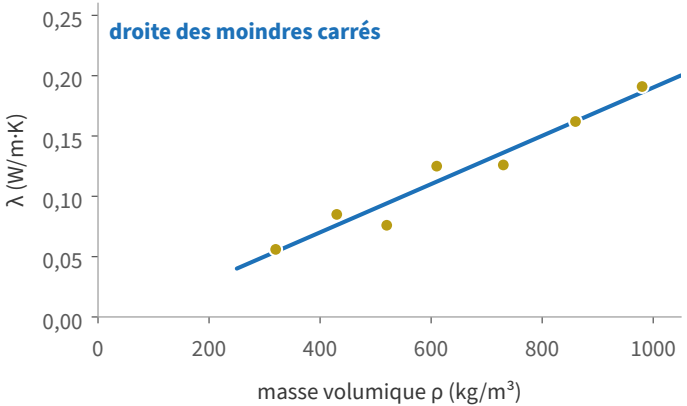
L’incertitude d’une droite

Sept formulations de terre allégée : la droite ajustée est elle-même incertaine.



L'incertitude d'une droite

Sept formulations de terre allégée : la droite ajustée est elle-même incertaine.



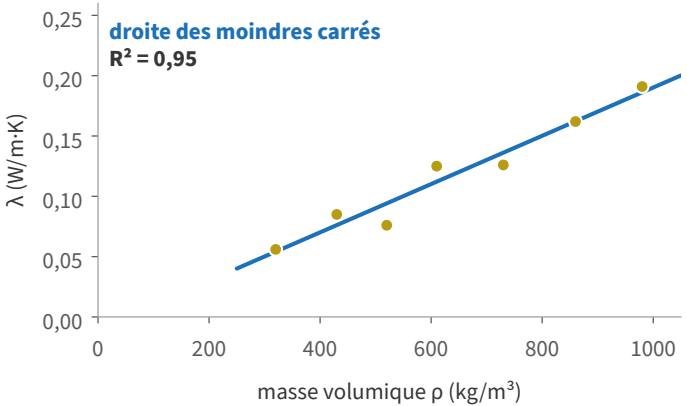
- Droite des moindres carrés.

Tableur : `=DROITEREG(plage_y ; plage_x ; VRAI ; VRAI)` (pente et son erreur standard)

Python : `r = stats.linregress(x, y); r.slope, stats.t.ppf(0.975, len(x)-2) * r.stderr`

L'incertitude d'une droite

Sept formulations de terre allégée : la droite ajustée est elle-même incertaine.



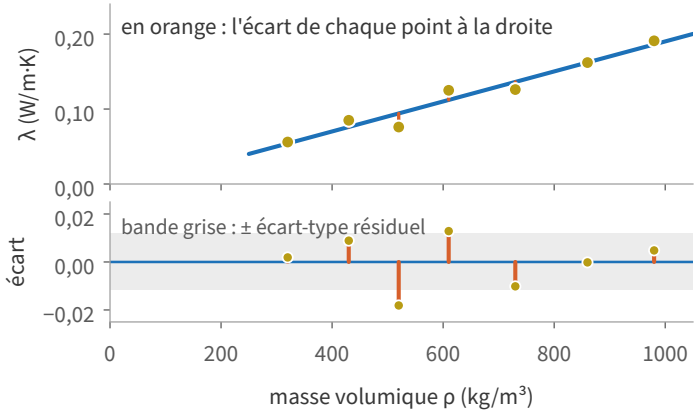
- Droite des moindres carrés.
- R^2 : part de la variabilité expliquée.

Tableur : `=DROITEREG(plage_y ; plage_x ; VRAI ; VRAI)` (pente et son erreur standard)

Python : `r = stats.linregress(x, y); r.slope, stats.t.ppf(0.975, len(x)-2) * r.stderr`

L'incertitude d'une droite

Sept formulations de terre allégée : la droite ajustée est elle-même incertaine.

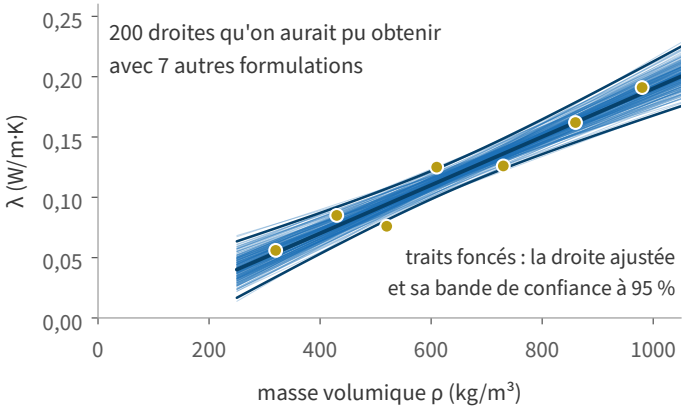


- Droite des moindres carrés.
- R^2 : part de la variabilité expliquée.
- Résidus : écarts à la droite. Normaux, indépendants, de même variance → moindres carrés = maximum de vraisemblance.

Tableur : `=DROITEREG(plage_y ; plage_x ; VRAI ; VRAI)` (pente et son erreur standard)
Python : `r = stats.linregress(x, y); r.slope, stats.t.ppf(0.975, len(x)-2) * r.stderr`

L'incertitude d'une droite

Sept formulations de terre allégée : la droite ajustée est elle-même incertaine.



- Droite des moindres carrés.
- R^2 : part de la variabilité expliquée.
- Résidus : écarts à la droite. Normaux, indépendants, de même variance → moindres carrés = maximum de vraisemblance.
- Pente : IC_{95} à 95 % de **± 26 %**.

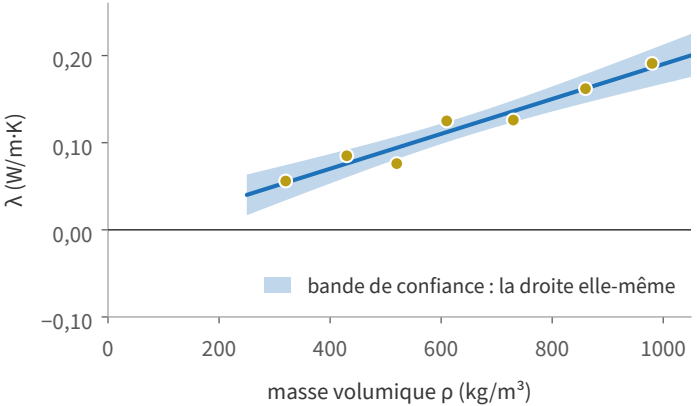
Tableur : `=DROITEREG(plage_y ; plage_x ; VRAI ; VRAI)` (pente et son erreur standard)
Python : `r = stats.linregress(x, y); r.slope, stats.t.ppf(0.975, len(x)-2) * r.stderr`

Confiance et prédiction

Confiance : où est la droite ; prédiction : où tombera une nouvelle formulation.

Confiance et prédiction

Confiance : où est la droite ; prédiction : où tombera une nouvelle formulation.

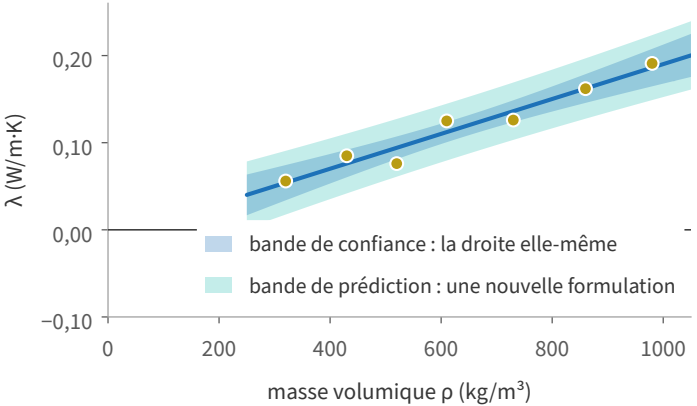


- **Bande de confiance** : où est la droite.

Tableur : les bandes demandent un calcul (pas de fonction directe)
Python : `sm.OLS(y, sm.add_constant(x)).fit().get_prediction([[1, 650]]).summary_frame()`

Confiance et prédiction

Confiance : où est la droite ; prédiction : où tombera une nouvelle formulation.



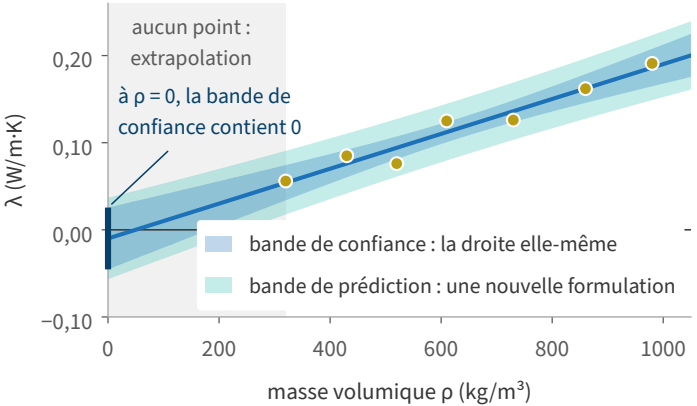
- **Bande de confiance** : où est la droite.
- **Bande de prédiction** : où tombera une nouvelle formulation.

Comme en partie 1 : confiance = IC_{95}, prédiction = IP_{95}.

Tableur : les bandes demandent un calcul (pas de fonction directe)
Python : `sm.OLS(y, sm.add_constant(x)).fit().get_prediction([[1, 650]]).summary_frame()`

Confiance et prédiction

Confiance : où est la droite ; prédiction : où tombera une nouvelle formulation.



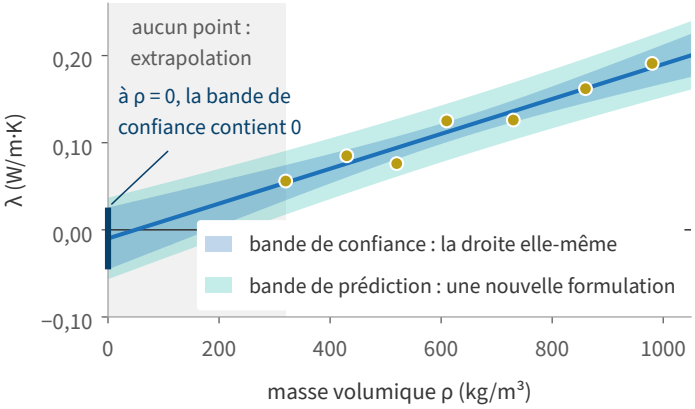
- **Bande de confiance** : où est la droite.
- **Bande de prédiction** : où tombera une nouvelle formulation.

Comme en partie 1 : confiance = IC_{95}, prédiction = IP_{95}.

- Ordonnée à l'origine : son IC_{95} contient 0 ; $\rho = 0$ est déjà une extrapolation.

Confiance et prédiction

Confiance : où est la droite ; prédiction : où tombera une nouvelle formulation.



- **Bande de confiance** : où est la droite.
- **Bande de prédiction** : où tombera une nouvelle formulation.

Comme en partie 1 : confiance = IC_{95}, prédiction = IP_{95}.

- Ordonnée à l'origine : son IC_{95} contient 0 ; $\rho = 0$ est déjà une extrapolation.

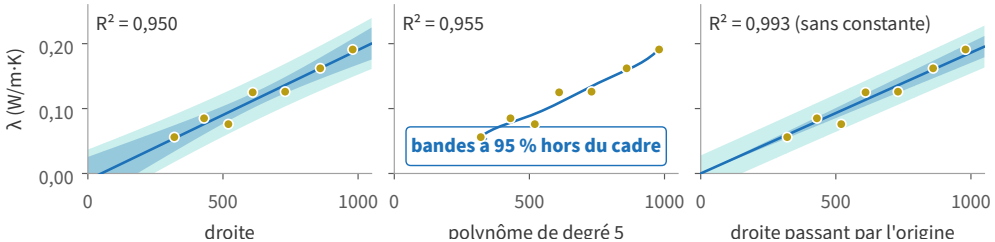
Cas publié : une cohésion lue à l'ordonnée à l'origine d'une droite ajustée, donnée sans incertitude.

Le R^2 ne suffit pas

Un R^2 élevé ne valide rien à lui seul.

Le R^2 ne suffit pas

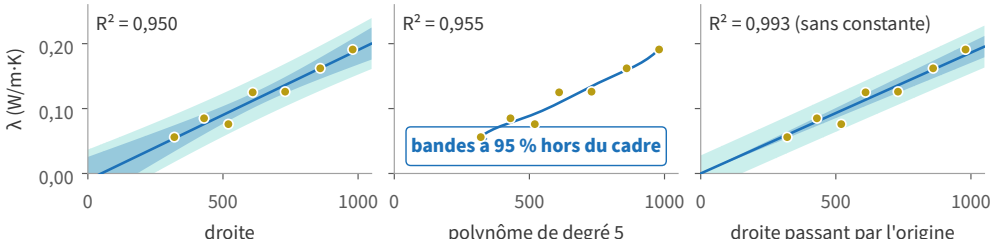
Un R^2 élevé ne valide rien à lui seul.



- Droite forcée par l'origine : R^2 sans constante, non comparable.

Le R^2 ne suffit pas

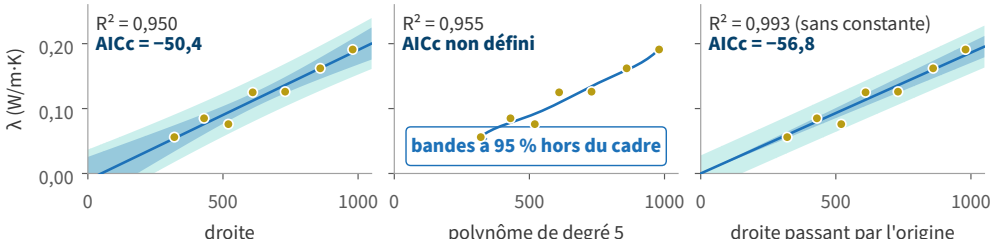
Un R^2 élevé ne valide rien à lui seul.



- Droite forcée par l'origine : R^2 sans constante, non comparable.
- Polynôme de degré 5 : 6 coefficients pour 7 points, bandes hors cadre.

Le R^2 ne suffit pas

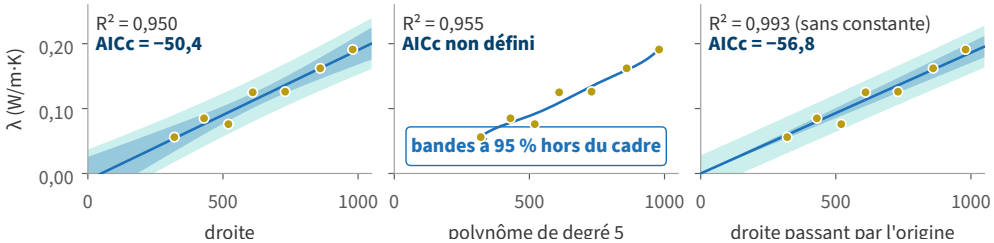
Un R^2 élevé ne valide rien à lui seul.



- Droite forcée par l'origine : R^2 sans constante, non comparable.
- Polynôme de degré 5 : 6 coefficients pour 7 points, bandes hors cadre.
- AICc : préfère la droite par l'origine (7 points : ordonnée non distinguable de 0, diapo 30)…

Le R^2 ne suffit pas

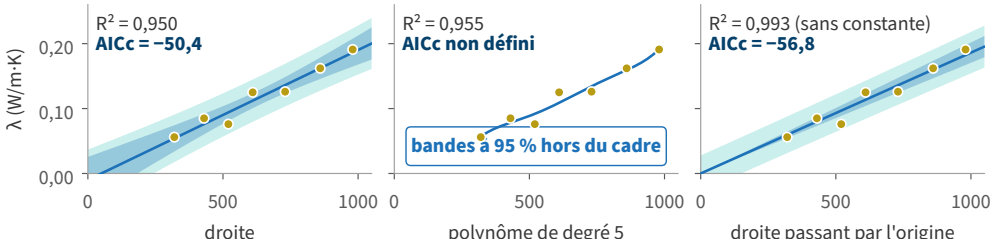
Un R^2 élevé ne valide rien à lui seul.



- Droite forcée par l'origine : R^2 sans constante, non comparable.
- Polynôme de degré 5 : 6 coefficients pour 7 points, bandes hors cadre.
- AICc : préfère la droite par l'origine (7 points : ordonnée non distinguable de 0, diapo 30)…
- **Physique** : $\lambda(0) = \lambda_{air} \approx 0{,}025$, ni −0,010 (libre) ni 0 (origine) → **aucune n'est physique en $\rho = 0$.**

Le R^2 ne suffit pas

Un R^2 élevé ne valide rien à lui seul.



- Droite forcée par l'origine : R^2 sans constante, non comparable.
- Polynôme de degré 5 : 6 coefficients pour 7 points, bandes hors cadre.
- AICc : préfère la droite par l'origine (7 points : ordonnée non distinguable de 0, diapo 30)…
- **Physique** : $\lambda(0) = \lambda_{air} \approx 0{,}025$, ni −0,010 (libre) ni 0 (origine) → **aucune n'est physique en $\rho = 0$.**

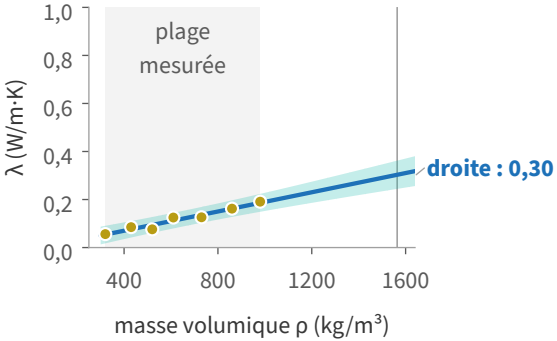
Avant de croire un R^2 : combien de points ? Quelle physique ? Que montrent les résidus et les bandes ?

Extrapoler

Hors des points mesurés, les modèles divergent.

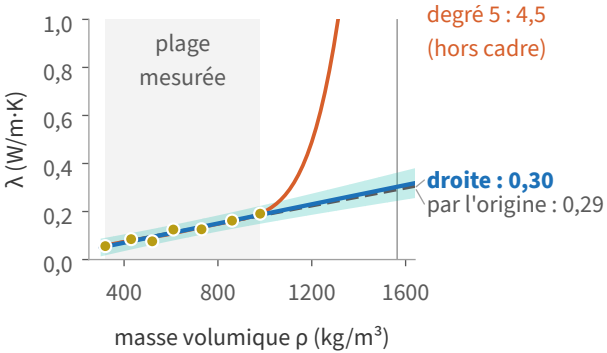
Extrapoler

Hors des points mesurés, les modèles divergent.



Extrapoler

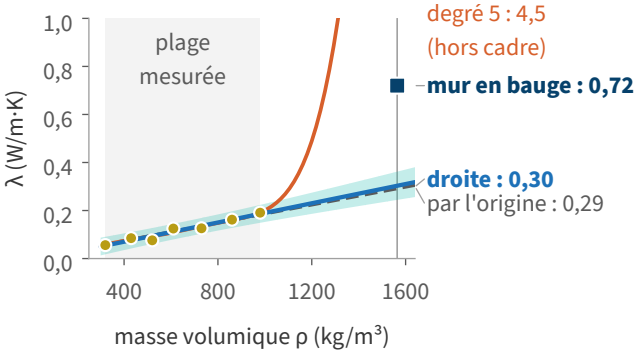
Hors des points mesurés, les modèles divergent.



- Hors de la plage mesurée, aucun critère (AICc, R^2) ne départage les modèles.

Extrapoler

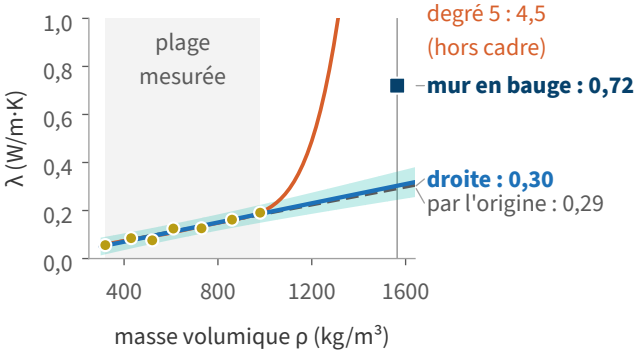
Hors des points mesurés, les modèles divergent.



- Hors de la plage mesurée, aucun critère (AICc, R^2) ne départage les modèles.
- L'IP_{95} suppose le modèle juste : il ne protège pas d'une extrapolation fausse.

Extrapoler

Hors des points mesurés, les modèles divergent.



- Hors de la plage mesurée, aucun critère (AICc, R^2) ne départage les modèles.
- L'IP_{95} suppose le modèle juste : il ne protège pas d'une extrapolation fausse.

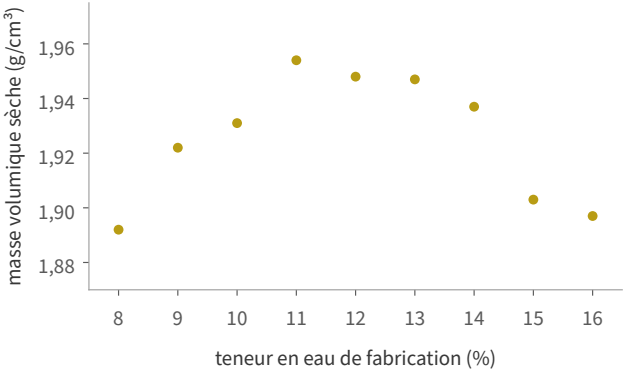
Cas publié : un modèle ajusté sur 4 teneurs, de 25 à 100 %, annoncé valable de 0 à 100 %.

Regarder avant d'ajuster

Une corrélation nulle n'exclut pas une relation forte.

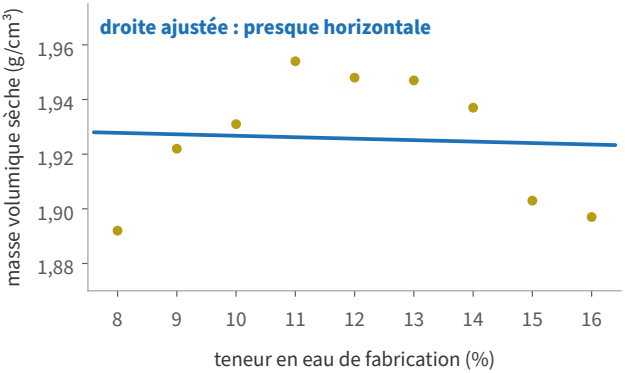
Regarder avant d'ajuster

Une corrélation nulle n'exclut pas une relation forte.



Regarder avant d'ajuster

Une corrélation nulle n'exclut pas une relation forte.

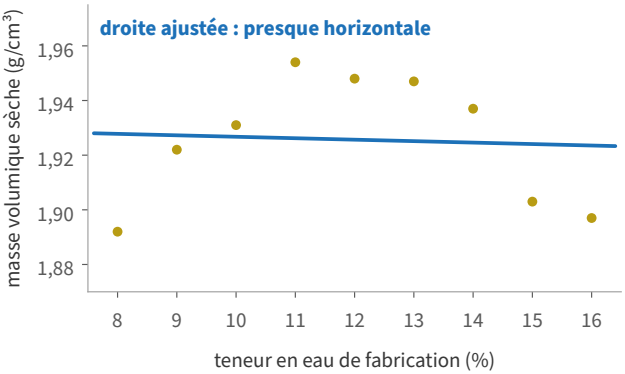


Coefficient de corrélation

- **Pearson** $r = -0{,}06$: relation **linéaire**.

Regarder avant d'ajuster

Une corrélation nulle n'exclut pas une relation forte.



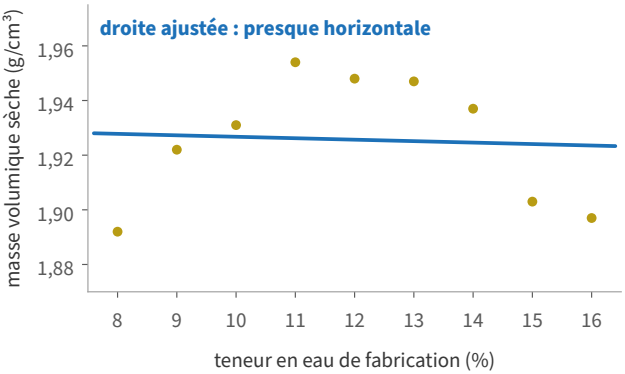
Coefficient de corrélation

- **Pearson** $r = -0{,}06$: relation **linéaire**.
- **Spearman** $r_S = +0{,}02$: relation monotone.
- ξ **de Chatterjee** $= 0{,}44$: toute dépendance, même en cloche.

Tableur : `=COEFFICIENT.CORRELATION(plage_x ; plage_y)`
Python : `stats.pearsonr(x, y)` ; `stats.spearmanr(x, y)` ; `stats.chatterjeexi(x, y)`

Regarder avant d'ajuster

Une corrélation nulle n'exclut pas une relation forte.



Coefficient de corrélation

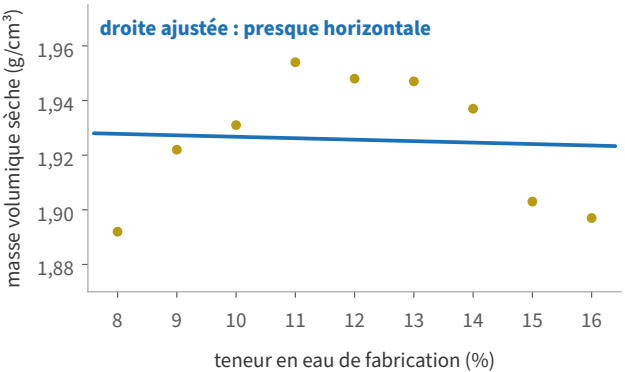
- **Pearson** $r = -0{,}06$: relation **linéaire**.
- **Spearman** $r_S = +0{,}02$: relation monotone.
- ξ **de Chatterjee** $= 0{,}44$: toute dépendance, même en cloche.

Un coefficient résume une forme de relation ; seul le graphique montre laquelle.

Tableur : `=COEFFICIENT.CORRELATION(plage_x ; plage_y)`
Python : `stats.pearsonr(x, y)` ; `stats.spearmanr(x, y)` ; `stats.chatterjeexi(x, y)`

Regarder avant d'ajuster

Une corrélation nulle n'exclut pas une relation forte.



Coefficient de corrélation

- **Pearson** $r = -0{,}06$: relation **linéaire**.
- **Spearman** $r_S = +0{,}02$: relation monotone.
- ξ **de Chatterjee** $= 0{,}44$: toute dépendance, même en cloche.

Un coefficient résume une forme de relation ; seul le graphique montre laquelle.

✔ Réponse 5

Que vaut mon ajustement ? **Tracer l'ajustement avec ses bandes, sans R^2 seul ni extrapolation ; une corrélation ne remplace pas le graphique.**

Objectif du cours : comprendre la variabilité, tester des hypothèses, planifier ses mesures

1. **Décrire la variabilité** Quelles notions pour décrire mes résultats ?
2. **Planifier** Combien d'éprouvettes ?
3. **Comparer** Deux échantillons sont-ils différents ?
4. **Choisir une loi** Quelle loi pour mes données ?
5. **Ajuster une tendance** Que vaut mon ajustement ?

   ✔ **Réponse** Tracer l'ajustement avec ses bandes, sans R^2 seul ni extrapolation ; une corrélation ne remplace pas le graphique.

6. **Les réflexes** **Quels réflexes garder ?**
7. **Ouverture** *Et après ?* *(après le quiz, si le temps le permet)*

Trois réflexes

Décider avant, regarder et écrire pendant, douter toujours.

Trois réflexes

Décider avant, regarder et écrire pendant, douter toujours.

Avant

de mesurer

Décider combien d'éprouvettes, quelle analyse, et évaluer l'incertitude de mesure (GUM).

Trois réflexes

Décider avant, regarder et écrire pendant, douter toujours.

Avant

de mesurer

Décider combien d'éprouvettes, quelle analyse, et évaluer l'incertitude de mesure (GUM).

Pendant

l'analyse

Regarder les données brutes ; écrire ce qu'on a fait.

Trois réflexes

Décider avant, regarder et écrire pendant, douter toujours.

Avant

de mesurer

Décider combien d'éprouvettes, quelle analyse, et évaluer l'incertitude de mesure (GUM).

Pendant

l'analyse

Regarder les données brutes ; écrire ce qu'on a fait.

Toujours

L'esprit critique, envers les résultats publiés, envers l'IA (exemple) et envers ses propres bases. D'où le besoin de se former.

Trois réflexes

Décider avant, regarder et écrire pendant, douter toujours.

Avant

de mesurer

Décider combien d'éprouvettes, quelle analyse, et évaluer l'incertitude de mesure (GUM).

Pendant

l'analyse

Regarder les données brutes ; écrire ce qu'on a fait.

Toujours

L'esprit critique, envers les résultats publiés, envers l'IA (exemple) et envers ses propres bases. D'où le besoin de se former.

Publié ne veut pas nécessairement dire vérifié sur le plan statistique.

Pour se former

Ce cours donne des repères ; ces ressources permettent de se former.

Concepts (introduction)
G. Baudin, *Estimation des incertitudes : cours et applications en langage Python*, Ellipses, 2020 (ISBN 978-2-340-03595-9).

Concepts (approfondir)
G. Saporta, *Probabilités, analyse des données et statistique*, Technip, 3e édition, 2011 (ISBN 978-2-7108-0980-7).

Incertitude de mesure
GUM (JCGM 100:2008), BIPM, gratuit : l'incertitude de mesure. Il ne dit ni comment caractériser la variabilité du matériau, ni combien d'éprouvettes prévoir, ni comment tester.
doi.org/10.59161/JCGM100-2008F

Cours en ligne
MIT OpenCourseWare (gratuits, en anglais) : *Introduction to Probability and Statistics* (18.05, 2022) ; *Statistical Thinking and Data Analysis* (15.075J, 2011).
ocw.mit.edu (18.05) ocw.mit.edu (15.075J)

Outils
Tableur ; Python (SciPy, statsmodels) ; entre les deux, **Jamovi** (libre).

Merci de votre attention

Statistiques pour expérimentateurs : comprendre la variabilité, tester des hypothèses et planifier ses mesures

Romain CLERC – GeM, UMR 6183, Nantes Université

Quiz Wooclap

Les messages du cours, mis en situation

Ouvrir le Wooclap

app.wooclap.com
code : **YLYNYCL**

Objectif du cours : comprendre la variabilité, tester des hypothèses, planifier ses mesures

1. **Décrire la variabilité** Quelles notions pour décrire mes résultats ?
2. **Planifier** Combien d'éprouvettes ?
3. **Comparer** Deux échantillons sont-ils différents ?
4. **Choisir une loi** Quelle loi pour mes données ?
5. **Ajuster une tendance** Que vaut mon ajustement ?
6. **Les réflexes** Quels réflexes garder ?

   ✔ **Réponse** Décider avant, regarder et écrire pendant, douter toujours ; se former pour le reste.

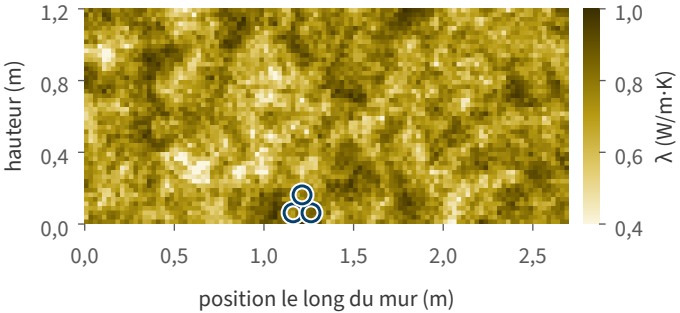
7. **Ouverture** **Et après ?**

Variabilité (et corrélation) spatiale

Des éprouvettes voisines se ressemblent : elles ne sont pas indépendantes.

Variabilité (et corrélation) spatiale

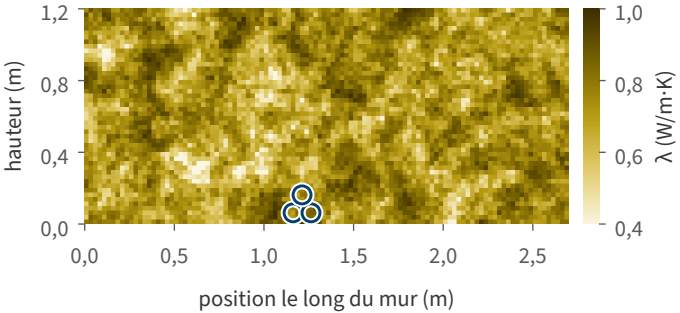
Des éprouvettes voisines se ressemblent : elles ne sont pas indépendantes.



- La statistique compresse l'information et se fonde sur des hypothèses à vérifier.

Variabilité (et corrélation) spatiale

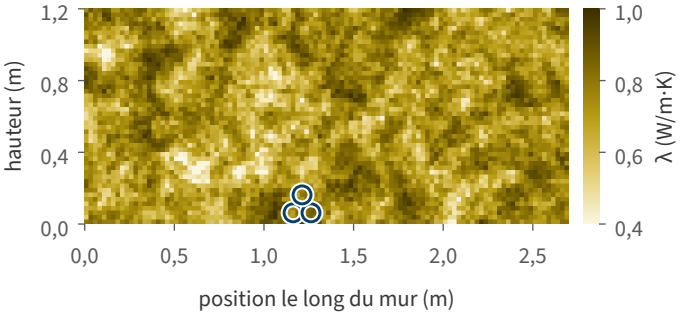
Des éprouvettes voisines se ressemblent : elles ne sont pas indépendantes.



- La statistique compresse l'information et se fonde sur des hypothèses à vérifier.
- Compte tenu de la corrélation entre voisins observée sur ce mur, trois éprouvettes jointives ne valent que **2 éprouvettes indépendantes** ; leur IC_{95} ne contient μ_λ que 92 fois sur 100.

Variabilité (et corrélation) spatiale

Des éprouvettes voisines se ressemblent : elles ne sont pas indépendantes.



- La statistique compresse l'information et se fonde sur des hypothèses à vérifier.
- Compte tenu de la corrélation entre voisins observée sur ce mur, trois éprouvettes jointives ne valent que **2 éprouvettes indépendantes** ; leur IC_{95} ne contient μ_λ que 92 fois sur 100.
- Idem pour une évolution temporelle : des mesures rapprochées dans le temps se ressemblent.

L’approche bayésienne : partir de ce qu’on sait

Trois éprouvettes disent peu seules, davantage si l’on ne part pas de zéro.

L’approche bayésienne : partir de ce qu’on sait

Trois éprouvettes disent peu seules, davantage si l’on ne part pas de zéro.

a posteriori $p(\mu_\lambda \mid \text{données})$

vraisemblance $\propto p(\text{données} \mid \mu_\lambda)$

a priori $\times\, p(\mu_\lambda)$

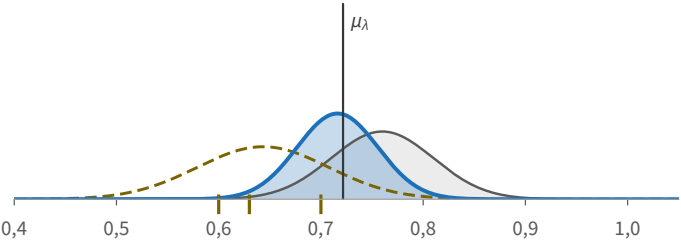
$p(A \mid B)$: probabilité de A sachant B (probabilité conditionnelle).

clic sur la figure : étape suivante (Acrobat Reader) folioscope (annexe A2)

L’approche bayésienne : partir de ce qu’on sait

Trois éprouvettes disent peu seules, davantage si l’on ne part pas de zéro.

a priori + 3 éprouvettes



μ_λ (W/m·K) ; traits sous l'axe : les éprouvettes mesurées

a posteriori $p(\mu_\lambda \mid \text{données})$

vraisemblance $\propto p(\text{données} \mid \mu_\lambda)$

a priori $\times\, p(\mu_\lambda)$

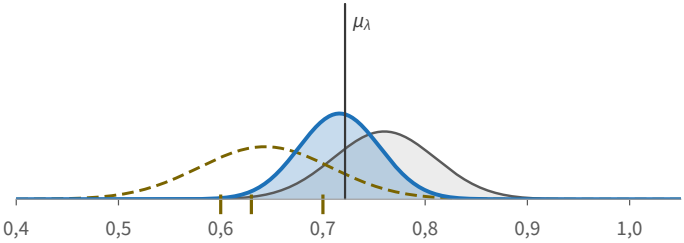
$p(A \mid B)$: probabilité de A sachant B (probabilité conditionnelle).

- A priori : les éprouvettes de laboratoire. A posteriori, avec les trois éprouvettes du début : un intervalle de crédibilité à 95 % plus étroit que l’IC_{95} sans a priori.
- Un a priori trop sûr impose sa valeur : il se justifie et s’écrit.

L'approche bayésienne : partir de ce qu'on sait

Trois éprouvettes disent peu seules, davantage si l'on ne part pas de zéro.

a priori + 3 éprouvettes



μ_λ (W/m·K) ; traits sous l'axe : les éprouvettes mesurées

a posteriori $p(\mu_\lambda \mid \text{données})$

vraisemblance $\propto p(\text{données} \mid \mu_\lambda)$

a priori $\times\ p(\mu_\lambda)$

$p(A \mid B)$: probabilité de A sachant B (probabilité conditionnelle).

✔ Réponse 7

Et après ? **Trois éprouvettes voisines ne valent pas trois éprouvettes indépendantes ; trois éprouvettes valent davantage avec ce qu'on savait déjà.**

Supplément – L'IA générative

Même question, deux réponses d'IA, la même assurance.

Question posée, mot pour mot : « lequel est le plus variable, et est-ce que la différence est significative ? » — bauge 0,839 ; 0,574 ; 0,707 ; béton de chanvre 0,084 ; 0,108 ; 0,100.

Claude « Haiku »

« La bauge (géosourcée) est nettement plus variable que le béton de chanvre, selon tous les critères […] F calculé = 117,4 ; F critique (α=0,05, ddl 2,2) = 19,0. Conclusion : La différence de variabilité est hautement significative ($p < 0,05$). »

Claude « Sonnet »

« Résumé : la bauge est plus variable en valeur absolue, mais pas en variabilité relative. Avec 3 éprouvettes, on ne peut pas conclure qu'un des deux matériaux est réellement plus variable. »

Essai du 07/10/2026 : deux modèles de Claude, dans l'environnement de travail de cette préparation (pas l'interface grand public) ; question posée telle quelle, une seule exécution par modèle ; extraits cités mot pour mot.

✔ **À vérifier soi-même** : moyennes × 7 → comparer les dispersions relatives (CV 19 % et 13 %) : test F sur les logarithmes, p = 0,63 → on ne peut pas conclure.

L'IA aide vraiment dès qu'on sait vérifier ce qu'elle produit.

Supplément – Vocabulaire français / anglais

Les termes du cours, tels qu'on les trouve dans les logiciels et les articles.

Français	Anglais
estimateur	*estimator*
justesse	*trueness*
fidélité	*precision*
exactitude	*accuracy*
erreur de mesure	*measurement error*
incertitude-type	*standard uncertainty*
incertitude de mesure	*measurement uncertainty*
écart-type	*standard deviation*
erreur standard	*standard error*
intervalle de confiance	*confidence interval*
intervalle de prédiction	*prediction interval*
coefficient de variation	*coefficient of variation*
unité de répétition	*experimental unit*
plan emboîté	*nested design*
hypothèse nulle	*null hypothesis*
statistique de test	*test statistic*
risque α	*type I error*

Français	Anglais
risque β	*type II error*
puissance	*power*
taille d'effet	*effect size*
vraisemblance	*likelihood*
AICc (Akaike corrigé)	*corrected Akaike information criterion*
moindres carrés	*least squares*
coefficient de détermination	*coefficient of determination*
résidus	*residuals*
bande de confiance	*confidence band*
bande de prédiction	*prediction band*
ordonnée à l'origine	*intercept*
coefficient de corrélation	*correlation coefficient*
ξ de Chatterjee	*Chatterjee's xi*
a priori, a posteriori	*prior, posterior*

S9 – Antisèche : tableur et Python (1/2)

Les calculs du cours, dans Excel (en français) et en Python.

Diapo	Calcul	Tableur	Python
8	écart-type, erreur standard	`=ECARTYPE.STANDARD(plage) ; =ECARTYPE.STANDARD(plage)/RACINE(NB(plage))`	`x.std(ddof=1) ; stats.sem(x)`
9	intervalle de confiance	`=INTERVALLE.CONFIANCE.STUDENT(0,05 ; ECARTYPE.STANDARD(plage) ; NB(plage))`	`stats.t.ppf(0.975, len(x)-1) * stats.sem(x)`
10	intervalle de prédiction	`=LOI.STUDENT.INVERSE.BILATERALE(0,05 ; NB(plage)-1) * ECARTYPE.STANDARD(plage) * RACINE(1+1/NB(plage))`	`stats.t.ppf(0.975, len(x)-1) * x.std(ddof=1) * np.sqrt(1 + 1/len(x))`
16	incertitude de mesure et matériau	utilitaire d'analyse, « Anova : un facteur » : $\hat{u}^2$ = CM intra ; $\hat{\sigma}^2_{\text{mat}}$ = (CM inter – $\hat{u}^2$)/r	`u2 = x.var(1, ddof=1).mean(); s2_mat = x.mean(1).var(ddof=1) - u2/x.shape[1]`
19	test de Welch	`=T.TEST(plage1 ; plage2 ; 2 ; 3)`	`stats.ttest_ind(a, b, equal_var=False)`

Noms des fonctions : Excel en français, vérifiés (site de Microsoft ou Excel).

S9 – Antisèche : tableur et Python (2/2)

Les calculs du cours, dans Excel (en français) et en Python.

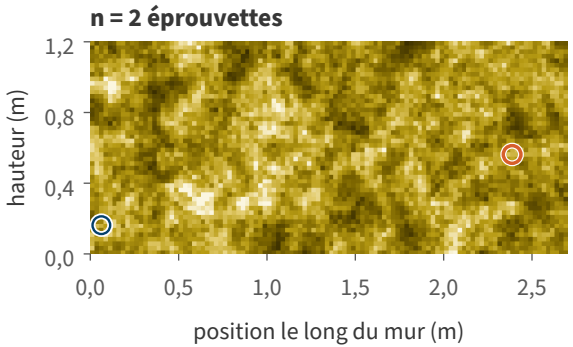
Diapo	Calcul	Tableur	Python
11, 41	CV; dispersions relatives	`=ECARTYPE.STANDARD(plage)/ MOYENNE(plage)` ; colonnes `=LN(x)`, puis `=F.TEST(plage_ln1 ; plage_ln2)`	`F = np.log(a).var(ddof=1) / np.log(b).var(ddof=1); d = len(a)-1, len(b)-1; 2*min(stats.f.sf(F, *d), stats.f.cdf(F, *d))`
29	pente et son incertitude	`=DROITEREG(plage_y ; plage_x ; VRAI ; VRAI)`	`r = stats.linregress(x, y); r.slope, stats.t.ppf(0.975, len(x)-2) * r.stderr`
30	bandes de confiance et de prédiction	pas de fonction directe	`sm.OLS(y, sm.add_constant(x)).fit() .get_prediction([[1, 650]]) .summary_frame(alpha=0.05)`
33	corrélation, dépendance	`=COEFFICIENT.CORRELATION(plage_x ; plage_y)` (pas de ξ)	`stats.pearsonr(x, y)` ; `stats.chatterjeexi(x, y)`

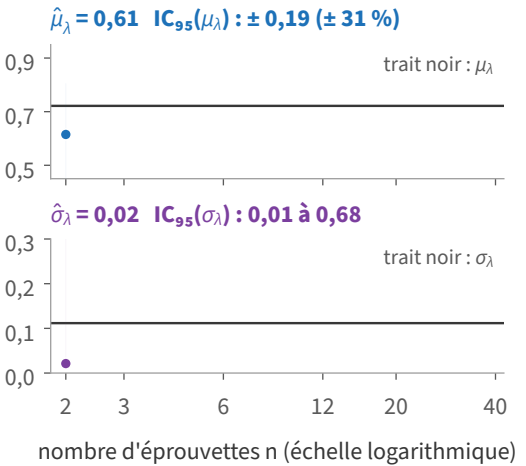
`import numpy as np ; from scipy import stats ; import statsmodels.api as sm` (SciPy 1.15 ou plus pour `chatterjeexi`)

Annexe A1 – La campagne qui grandit

Une image par clic : $n = 2$ à 40, puis l'image finale.

retour à la diapo 14

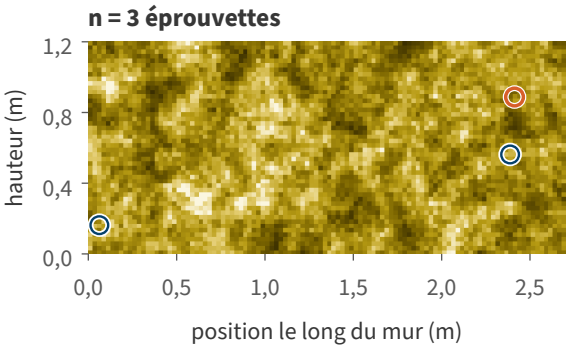


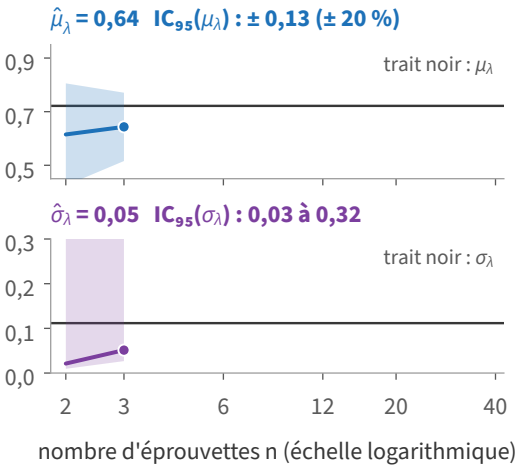


Annexe A1 – La campagne qui grandit

Une image par clic : $n = 2$ à 40, puis l'image finale.

retour à la diapo 14

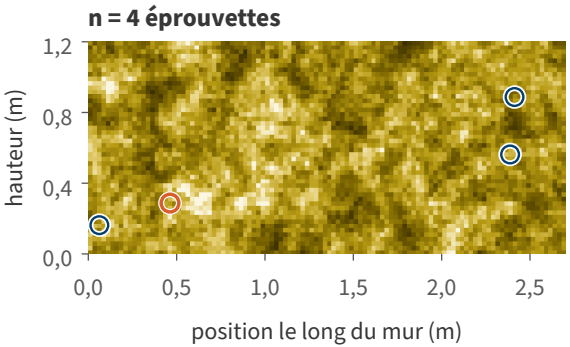


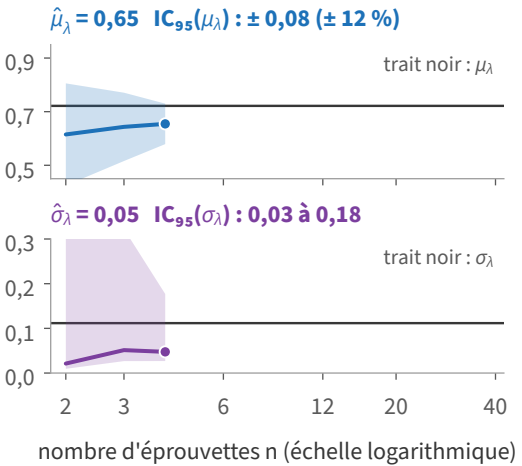


Annexe A1 – La campagne qui grandit

Une image par clic : $n = 2$ à 40, puis l'image finale.

retour à la diapo 14

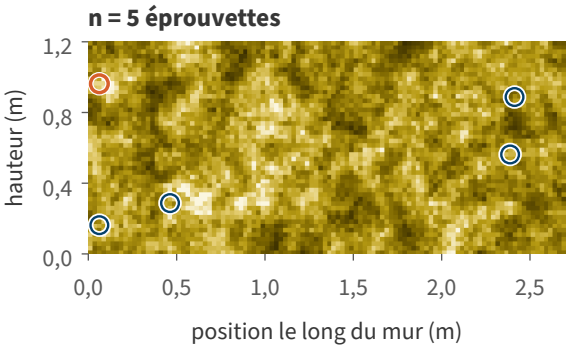


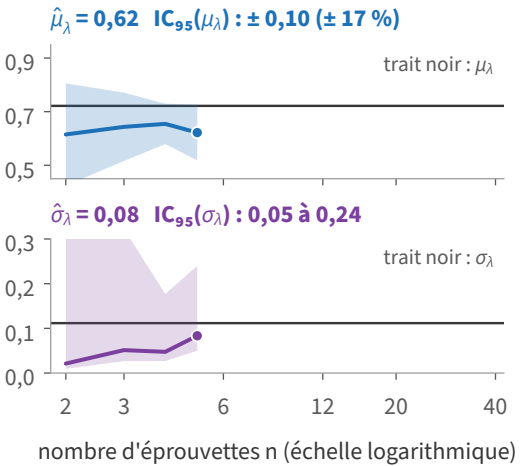


Annexe A1 – La campagne qui grandit

Une image par clic : $n = 2$ à 40, puis l'image finale.

retour à la diapo 14

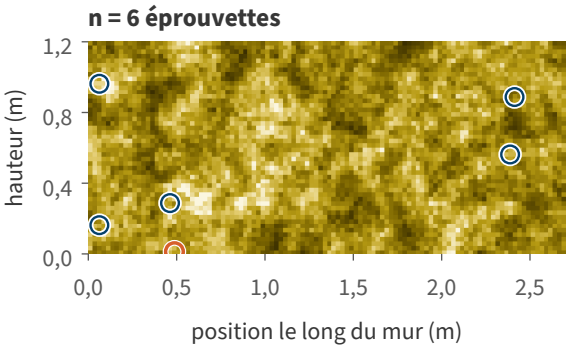


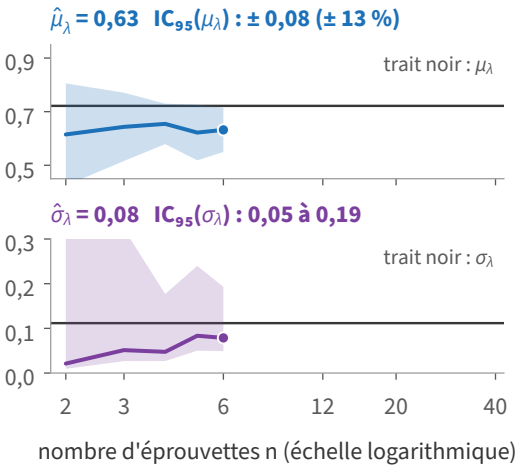


Annexe A1 – La campagne qui grandit

Une image par clic : $n = 2$ à 40, puis l'image finale.

retour à la diapo 14

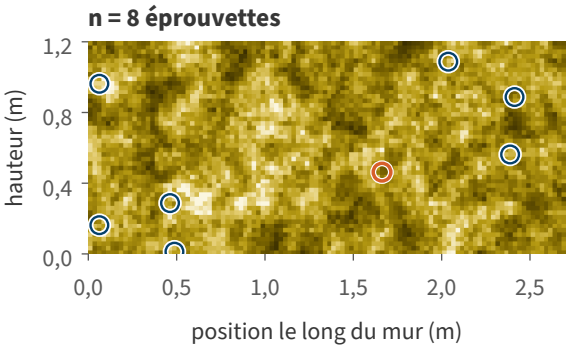


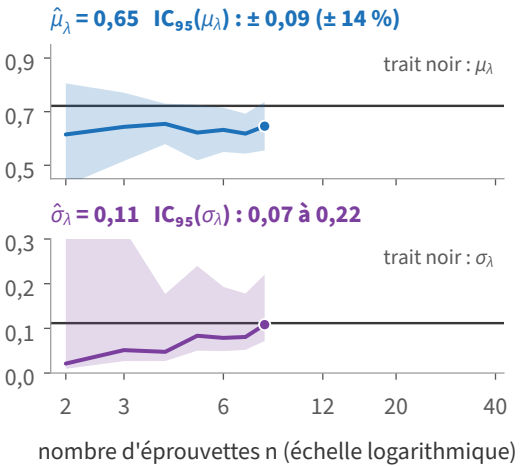


Annexe A1 – La campagne qui grandit

Une image par clic : $n = 2$ à 40, puis l'image finale.

retour à la diapo 14

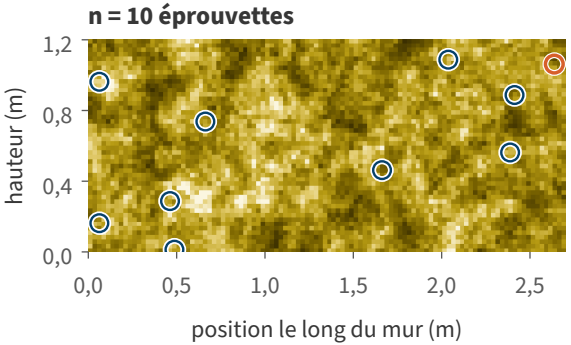


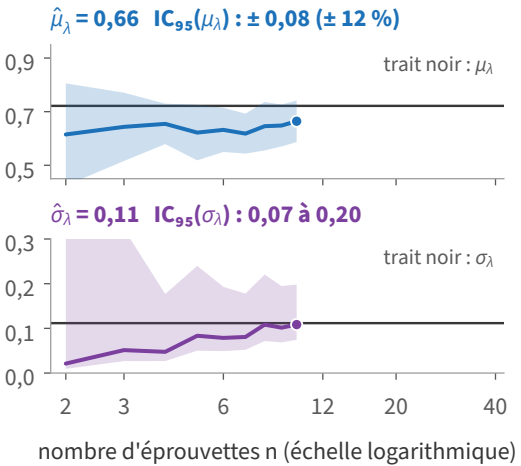


Annexe A1 – La campagne qui grandit

Une image par clic : $n = 2$ à 40, puis l'image finale.

retour à la diapo 14

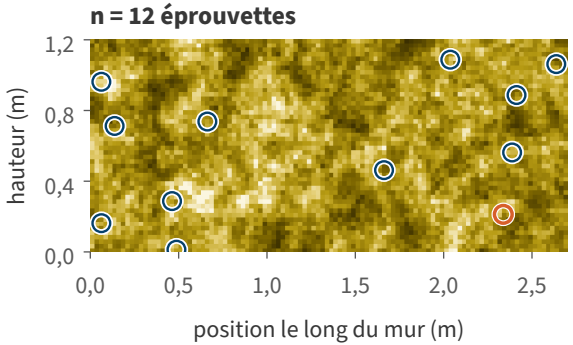


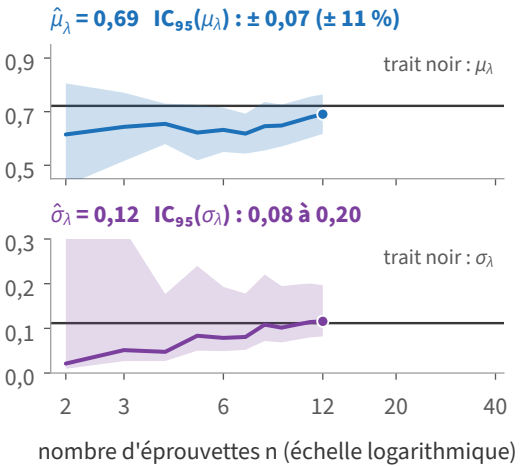


Annexe A1 – La campagne qui grandit

Une image par clic : $n = 2$ à 40, puis l'image finale.

retour à la diapo 14

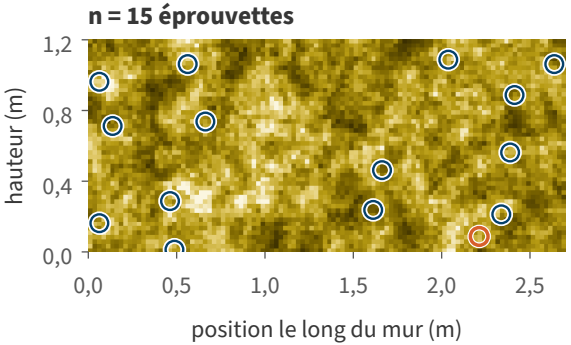


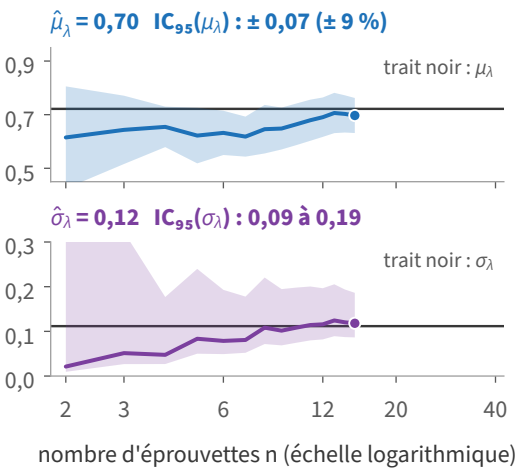


Annexe A1 – La campagne qui grandit

Une image par clic : $n = 2$ à 40, puis l'image finale.

retour à la diapo 14

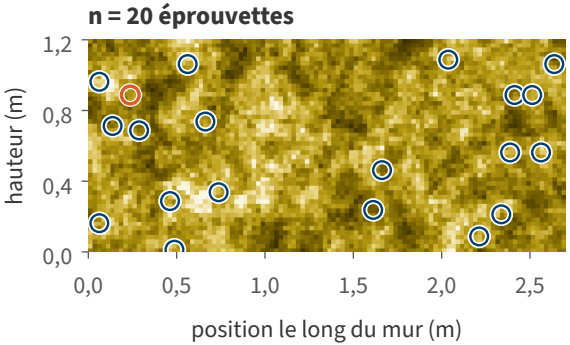


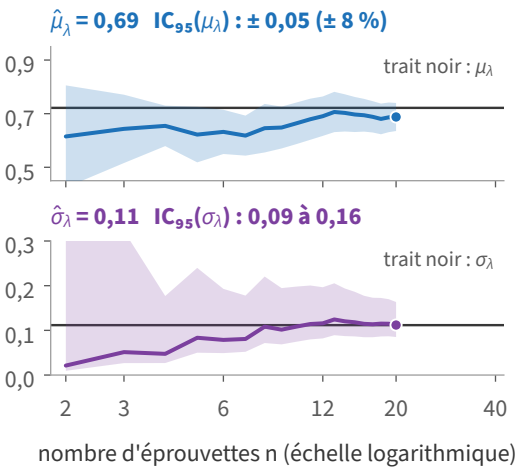


Annexe A1 – La campagne qui grandit

Une image par clic : $n = 2$ à 40, puis l'image finale.

retour à la diapo 14

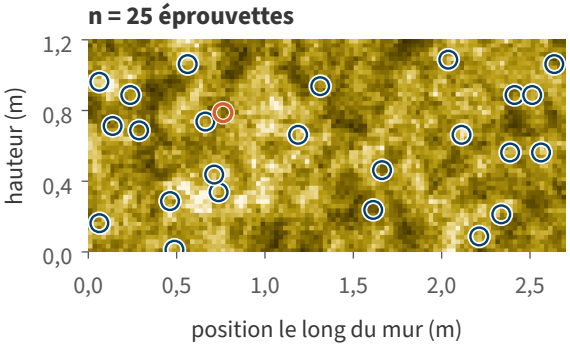


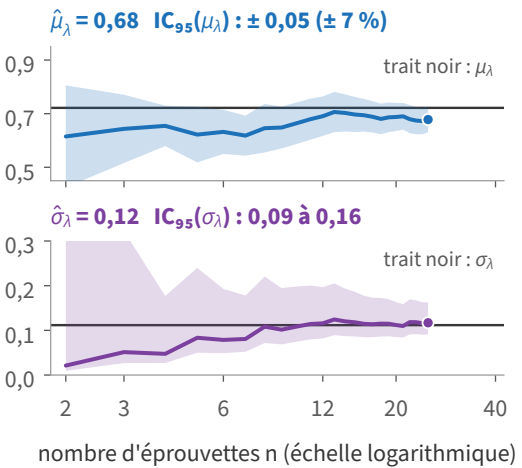


Annexe A1 – La campagne qui grandit

Une image par clic : $n = 2$ à 40, puis l'image finale.

retour à la diapo 14

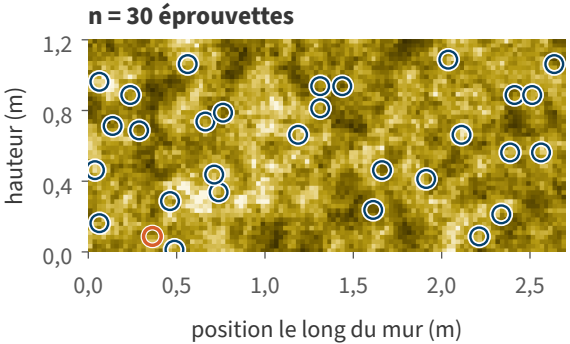


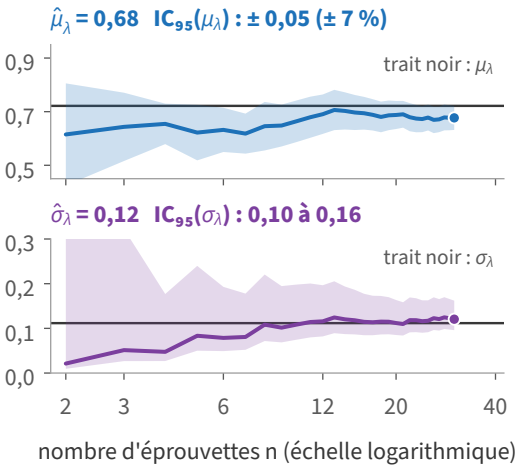


Annexe A1 – La campagne qui grandit

Une image par clic : $n = 2$ à 40, puis l'image finale.

retour à la diapo 14

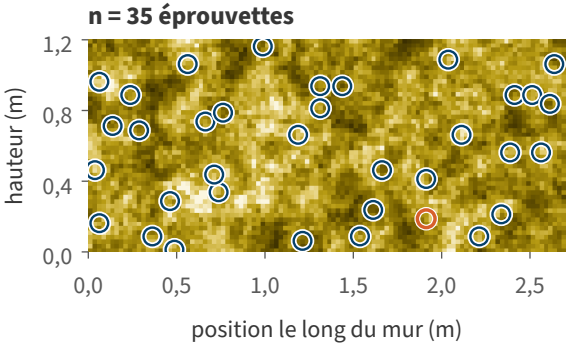


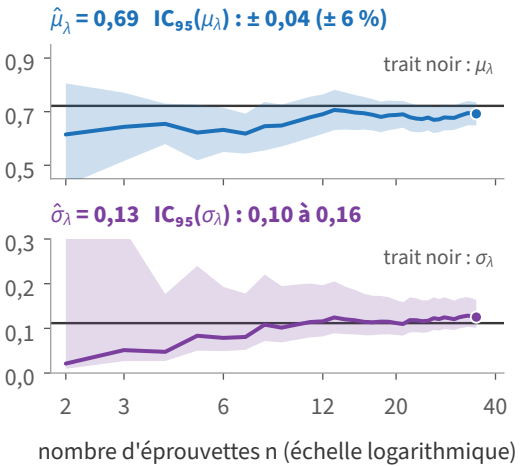


Annexe A1 – La campagne qui grandit

Une image par clic : $n = 2$ à 40, puis l'image finale.

retour à la diapo 14

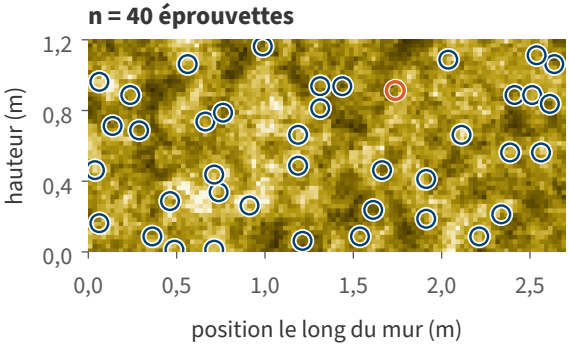


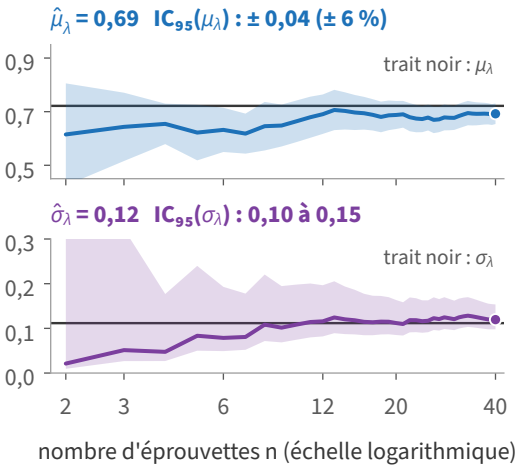


Annexe A1 – La campagne qui grandit

Une image par clic : $n = 2$ à 40, puis l'image finale.

retour à la diapo 14

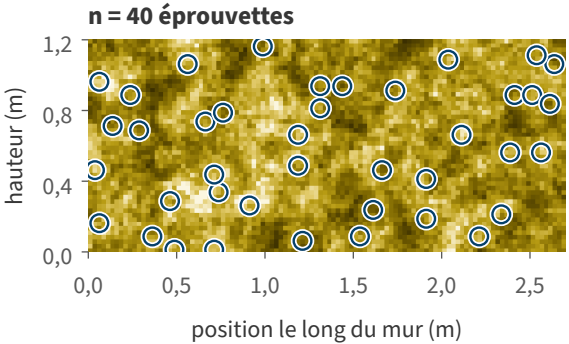


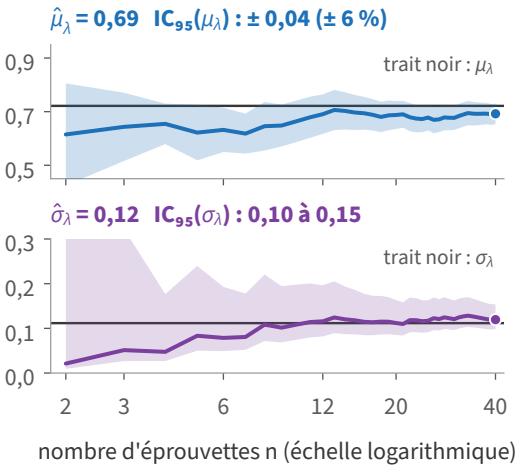


Annexe A1 – La campagne qui grandit

Une image par clic : $n = 2$ à 40, puis l'image finale.

retour à la diapo 14



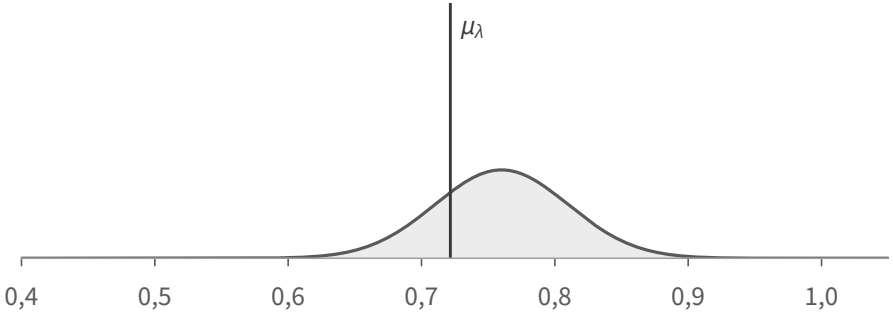


Annexe A2 – L’approche bayésienne

Une image par clic : l’a priori seul, puis 1, 2, 3, 6 et 12 éprouvettes.

retour à la diapo 40

a priori seul



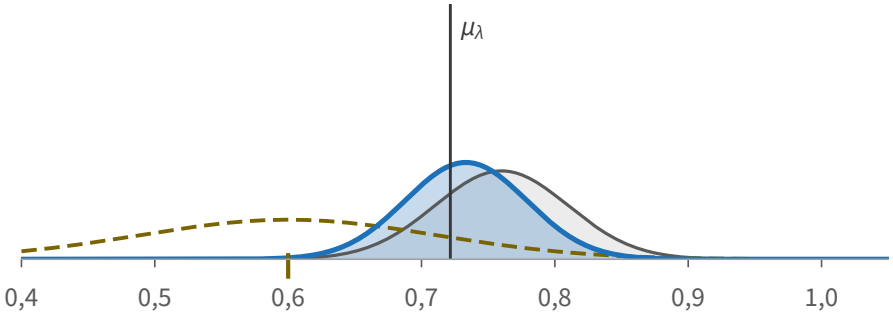
μ_λ (W/m·K) ; traits sous l'axe : les éprouvettes mesurées

Annexe A2 – L'approche bayésienne

Une image par clic : l'a priori seul, puis 1, 2, 3, 6 et 12 éprouvettes.

retour à la diapo 40

a priori + 1 éprouvette



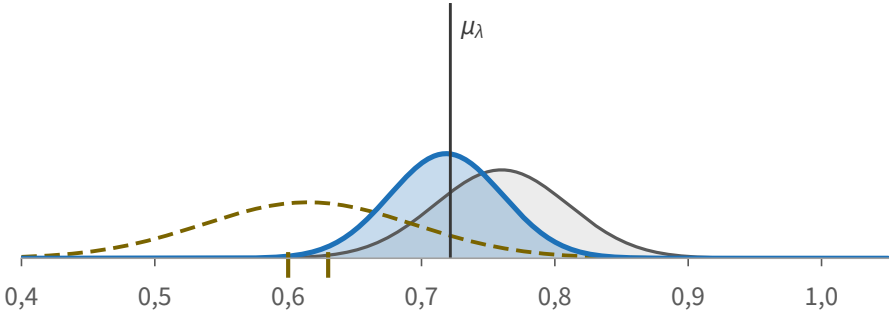
μ_λ (W/m·K) ; traits sous l'axe : les éprouvettes mesurées

Annexe A2 – L'approche bayésienne

Une image par clic : l'a priori seul, puis 1, 2, 3, 6 et 12 éprouvettes.

retour à la diapo 40

a priori + 2 éprouvettes



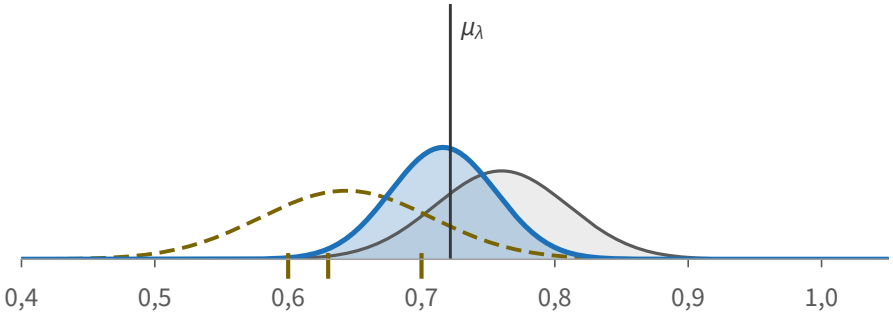
μ_λ (W/m·K) ; traits sous l'axe : les éprouvettes mesurées

Annexe A2 – L'approche bayésienne

Une image par clic : l'a priori seul, puis 1, 2, 3, 6 et 12 éprouvettes.

retour à la diapo 40

a priori + 3 éprouvettes



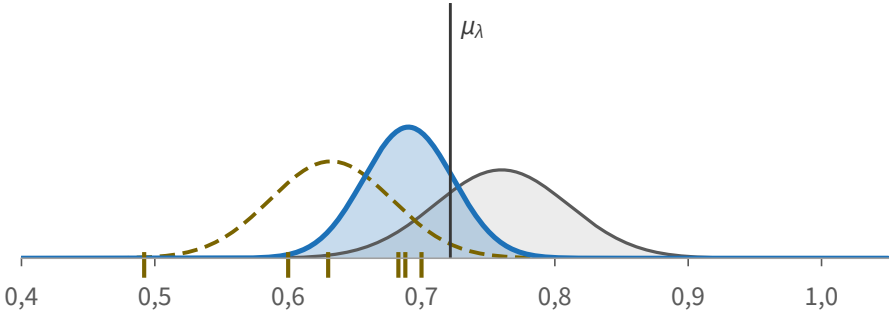
μ_λ (W/m·K) ; traits sous l'axe : les éprouvettes mesurées

Annexe A2 – L'approche bayésienne

Une image par clic : l'a priori seul, puis 1, 2, 3, 6 et 12 éprouvettes.

retour à la diapo 40

a priori + 6 éprouvettes



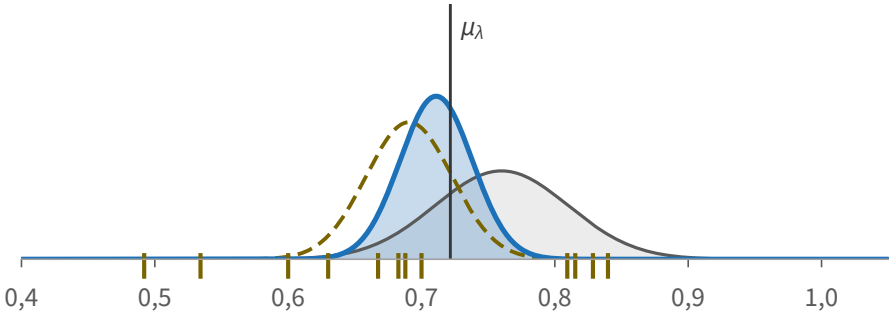
μ_λ (W/m·K) ; traits sous l'axe : les éprouvettes mesurées

Annexe A2 – L'approche bayésienne

Une image par clic : l'a priori seul, puis 1, 2, 3, 6 et 12 éprouvettes.

retour à la diapo 40

a priori + 12 éprouvettes

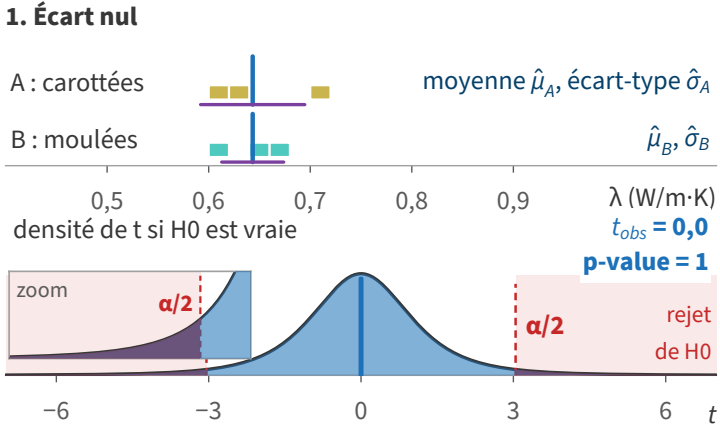


μ_λ (W/m·K) ; traits sous l'axe : les éprouvettes mesurées

Annexe A3 – Les cinq cas de la diapo 20

Une image par clic ; retour à la diapo 20 par le lien en haut à droite.

retour à la diapo 20

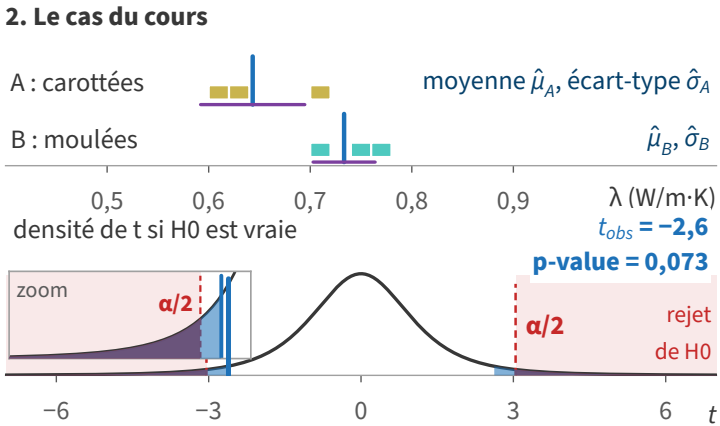


Cas construits à partir des valeurs de la diapo 19 (écart ou dispersion modifiés), pas tirés du mur.

Annexe A3 – Les cinq cas de la diapo 20

Une image par clic ; retour à la diapo 20 par le lien en haut à droite.

retour à la diapo 20

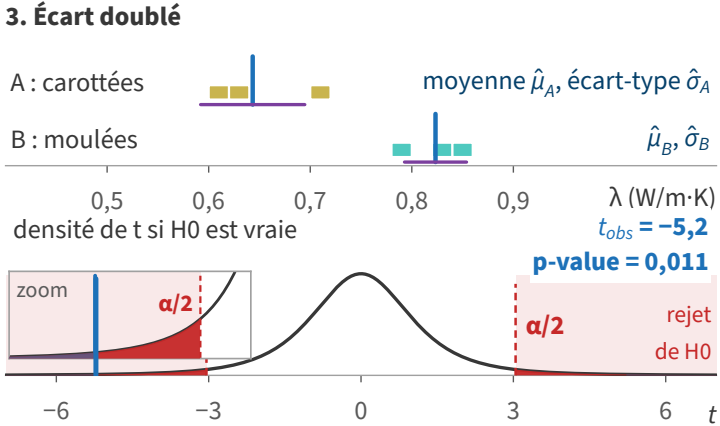


Cas construits à partir des valeurs de la diapo 19 (écart ou dispersion modifiés), pas tirés du mur.

Annexe A3 – Les cinq cas de la diapo 20

Une image par clic ; retour à la diapo 20 par le lien en haut à droite.

retour à la diapo 20

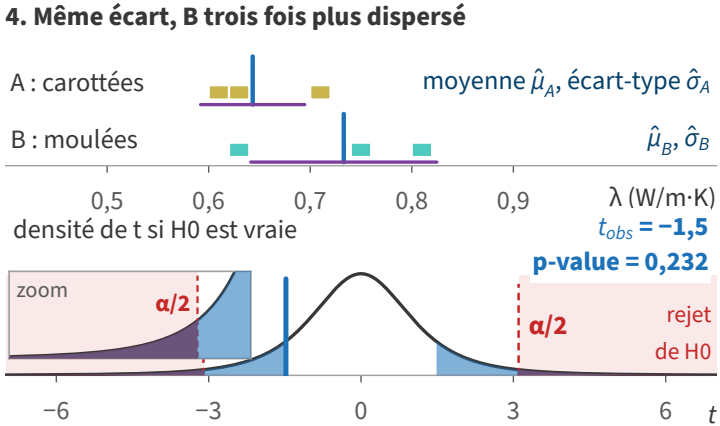


Cas construits à partir des valeurs de la diapo 19 (écart ou dispersion modifiés), pas tirés du mur.

Annexe A3 – Les cinq cas de la diapo 20

Une image par clic ; retour à la diapo 20 par le lien en haut à droite.

retour à la diapo 20



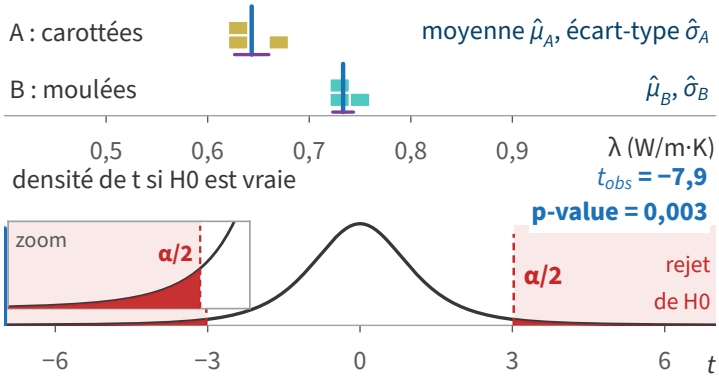
Cas construits à partir des valeurs de la diapo 19 (écart ou dispersion modifiés), pas tirés du mur.

Annexe A3 – Les cinq cas de la diapo 20

Une image par clic ; retour à la diapo 20 par le lien en haut à droite.

retour à la diapo 20

5. Même écart, échantillons trois fois moins dispersés

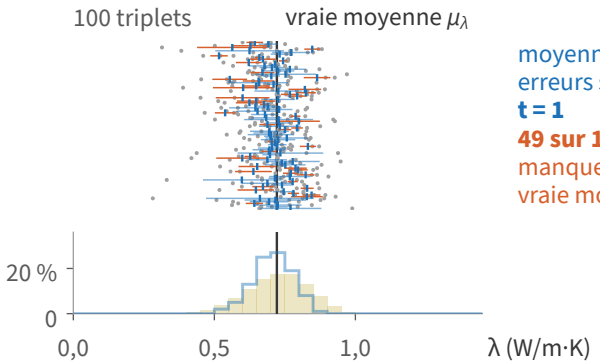


Cas construits à partir des valeurs de la diapo 19 (écart ou dispersion modifiés), pas tirés du mur.

Annexe A4 – L'intervalle $\hat{\mu}_\lambda \pm t \times \hat{\sigma}_{\hat{\mu}_\lambda}$ quand t grandit

Une image par clic : t = 1 ; 1,5 ; 2 ; 2,5 ; 3 ; 3,5 ; 4,30 ; 5 ; 6.

retour à la diapo 9



moyenne ± t erreurs standard

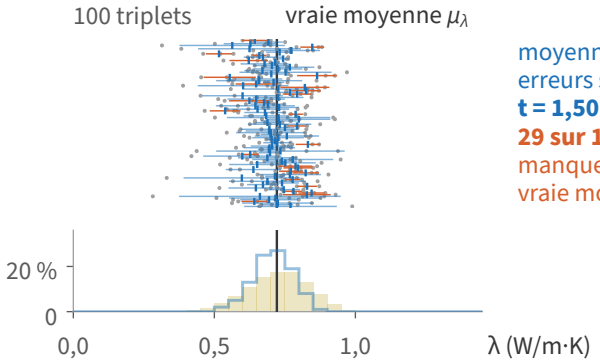
t = 1

49 sur 100 manquent la vraie moyenne

Annexe A4 – L'intervalle $\hat{\mu}_\lambda \pm t \times \hat{\sigma}_{\hat{\mu}_\lambda}$ quand t grandit

Une image par clic : t = 1 ; 1,5 ; 2 ; 2,5 ; 3 ; 3,5 ; 4,30 ; 5 ; 6.

retour à la diapo 9

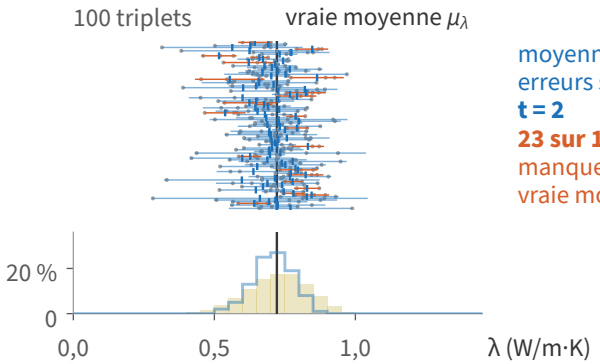


moyenne ± t erreurs standard
t = 1,50
29 sur 100 manquent la vraie moyenne

Annexe A4 – L'intervalle $\hat{\mu}_\lambda \pm t \times \hat{\sigma}_{\hat{\mu}_\lambda}$ quand t grandit

Une image par clic : t = 1 ; 1,5 ; 2 ; 2,5 ; 3 ; 3,5 ; 4,30 ; 5 ; 6.

retour à la diapo 9



moyenne ± t erreurs standard

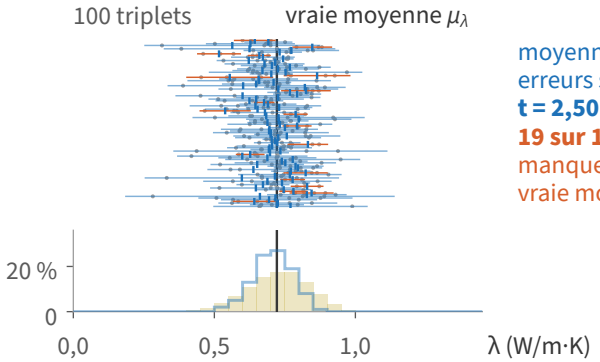
t = 2

23 sur 100 manquent la vraie moyenne

Annexe A4 – L'intervalle $\hat{\mu}_\lambda \pm t \times \hat{\sigma}_{\hat{\mu}_\lambda}$ quand t grandit

Une image par clic : t = 1 ; 1,5 ; 2 ; 2,5 ; 3 ; 3,5 ; 4,30 ; 5 ; 6.

retour à la diapo 9



moyenne ± t erreurs standard

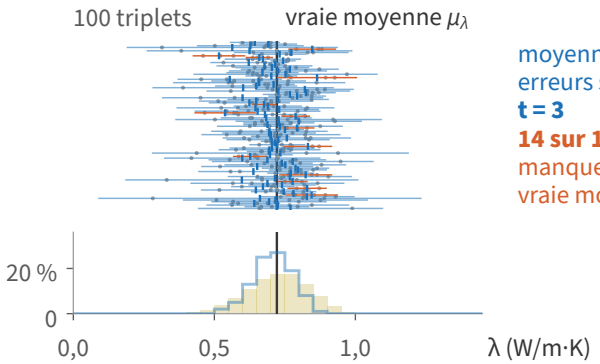
t = 2,50

19 sur 100 manquent la vraie moyenne

Annexe A4 – L'intervalle $\hat{\mu}_\lambda \pm t \times \hat{\sigma}_{\hat{\mu}_\lambda}$ quand t grandit

Une image par clic : t = 1 ; 1,5 ; 2 ; 2,5 ; 3 ; 3,5 ; 4,30 ; 5 ; 6.

retour à la diapo 9



moyenne ± t erreurs standard

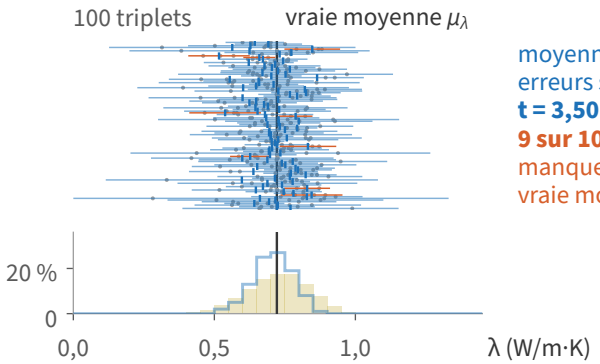
t = 3

14 sur 100 manquent la vraie moyenne

Annexe A4 – L'intervalle $\hat{\mu}_\lambda \pm t \times \hat{\sigma}_{\hat{\mu}_\lambda}$ quand t grandit

Une image par clic : t = 1 ; 1,5 ; 2 ; 2,5 ; 3 ; 3,5 ; 4,30 ; 5 ; 6.

retour à la diapo 9

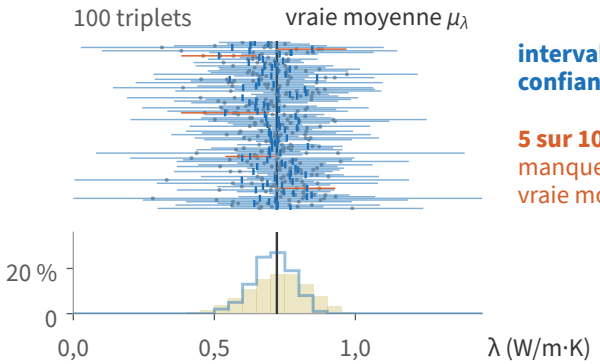


moyenne ± t erreurs standard
t = 3,50
9 sur 100 manquent la vraie moyenne

Annexe A4 – L'intervalle $\hat{\mu}_\lambda \pm t \times \hat{\sigma}_{\hat{\mu}_\lambda}$ quand t grandit

Une image par clic : t = 1 ; 1,5 ; 2 ; 2,5 ; 3 ; 3,5 ; 4,30 ; 5 ; 6.

retour à la diapo 9



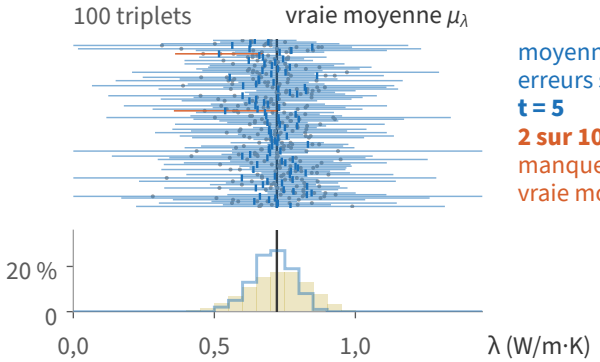
intervalle de confiance à 95 %

5 sur 100 manquent la vraie moyenne

Annexe A4 – L'intervalle $\hat{\mu}_\lambda \pm t \times \hat{\sigma}_{\hat{\mu}_\lambda}$ quand t grandit

Une image par clic : t = 1 ; 1,5 ; 2 ; 2,5 ; 3 ; 3,5 ; 4,30 ; 5 ; 6.

retour à la diapo 9

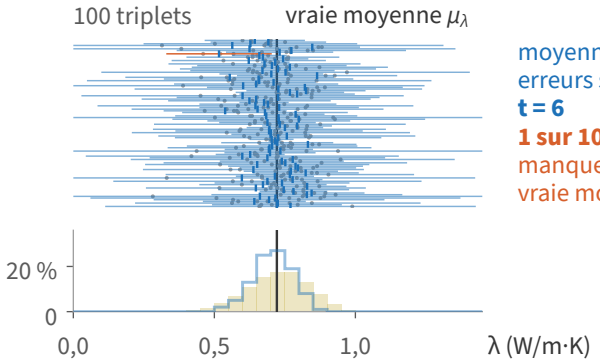


moyenne ± t erreurs standard
t = 5
2 sur 100 manquent la vraie moyenne

Annexe A4 – L'intervalle $\hat{\mu}_\lambda \pm t \times \hat{\sigma}_{\hat{\mu}_\lambda}$ quand t grandit

Une image par clic : t = 1 ; 1,5 ; 2 ; 2,5 ; 3 ; 3,5 ; 4,30 ; 5 ; 6.

retour à la diapo 9



moyenne ± t erreurs standard

t = 6

1 sur 100 manquent la vraie moyenne

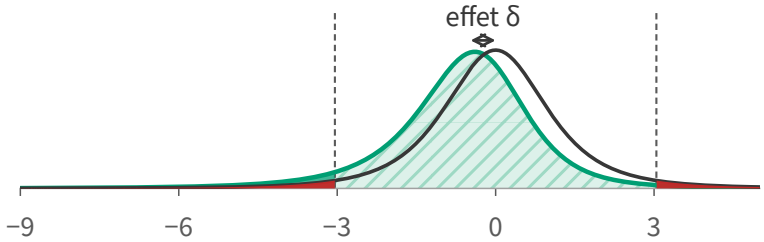
Annexe A5 – Puissance, effet et nombre d'éprouvettes

Une image par clic : $n = 3$ à 105 à effet fixe, puis effets plus grands à $n = 3$.

retour à la diapo 21

n = 3 éprouvettes par échantillon **puissance 1 − β = 6,5 %**

taille d'effet δ/σ = 0,4



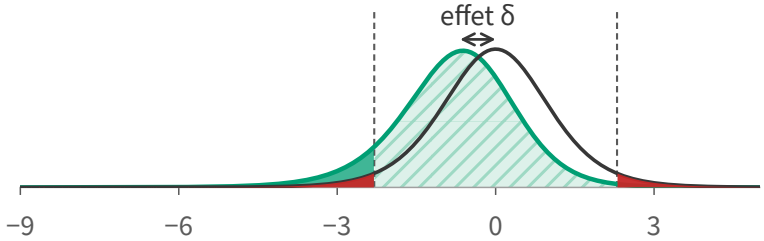
Annexe A5 – Puissance, effet et nombre d'éprouvettes

Une image par clic : $n = 3$ à 105 à effet fixe, puis effets plus grands à $n = 3$.

retour à la diapo 21

n = 6 éprouvettes par échantillon **puissance 1 − β = 9,2 %**

taille d'effet δ/σ = 0,4



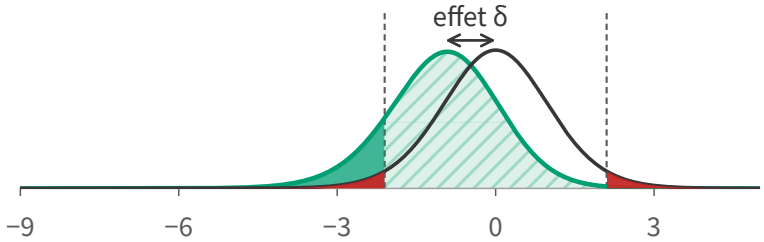
Annexe A5 – Puissance, effet et nombre d'éprouvettes

Une image par clic : $n = 3$ à 105 à effet fixe, puis effets plus grands à $n = 3$.

retour à la diapo 21

n = 12 éprouvettes par échantillon **puissance 1 − β = 14,7 %**

taille d'effet δ/σ = 0,4



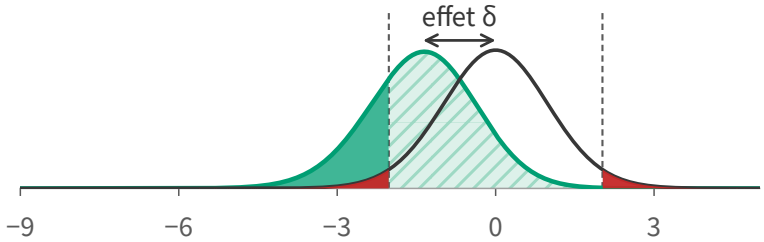
Annexe A5 – Puissance, effet et nombre d'éprouvettes

Une image par clic : $n = 3$ à 105 à effet fixe, puis effets plus grands à $n = 3$.

retour à la diapo 21

n = 25 éprouvettes par échantillon **puissance 1 − β = 26,8 %**

taille d'effet δ/σ = 0,4



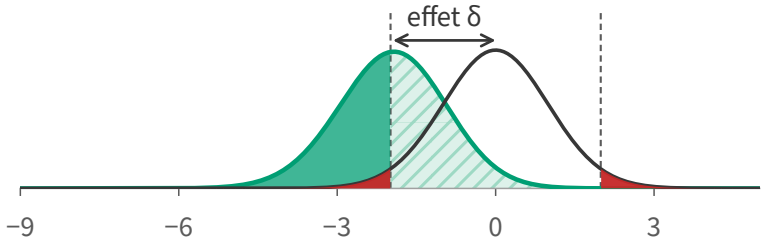
Annexe A5 – Puissance, effet et nombre d'éprouvettes

Une image par clic : $n = 3$ à 105 à effet fixe, puis effets plus grands à $n = 3$.

retour à la diapo 21

n = 50 éprouvettes par échantillon **puissance 1 − β = 48,4 %**

taille d'effet δ/σ = 0,4



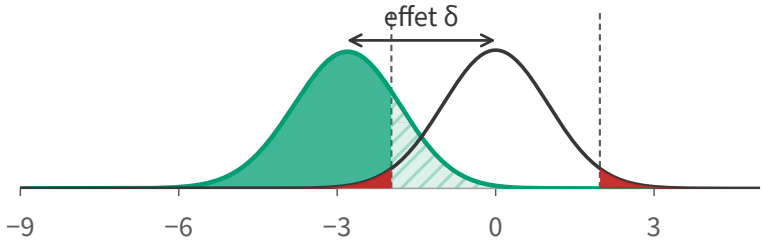
Annexe A5 – Puissance, effet et nombre d'éprouvettes

Une image par clic : $n = 3$ à 105 à effet fixe, puis effets plus grands à $n = 3$.

retour à la diapo 21

n = 105 éprouvettes par échantillon **puissance 1 − β = 80,0 %**

taille d'effet δ/σ = 0,4



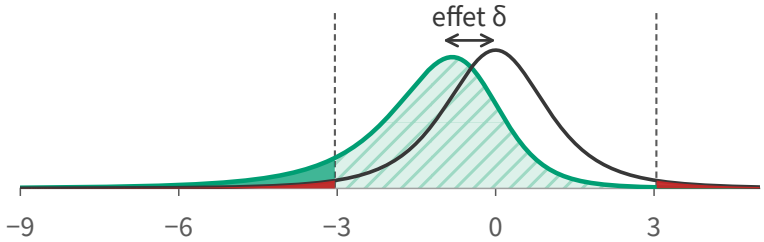
Annexe A5 – Puissance, effet et nombre d'éprouvettes

Une image par clic : $n = 3$ à 105 à effet fixe, puis effets plus grands à $n = 3$.

retour à la diapo 21

n = 3 éprouvettes par échantillon **puissance 1 − β = 11,4 %**

taille d'effet δ/σ = 0,8



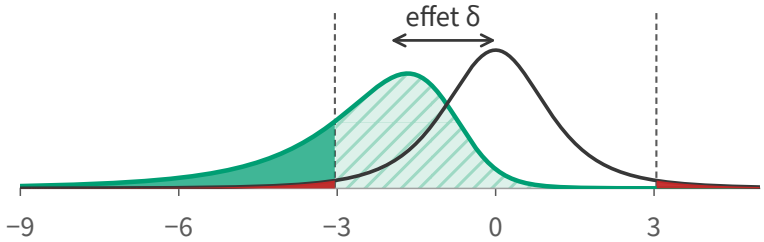
Annexe A5 – Puissance, effet et nombre d'éprouvettes

Une image par clic : $n = 3$ à 105 à effet fixe, puis effets plus grands à $n = 3$.

retour à la diapo 21

n = 3 éprouvettes par échantillon **puissance 1 − β = 30,0 %**

taille d'effet δ/σ = 1,6



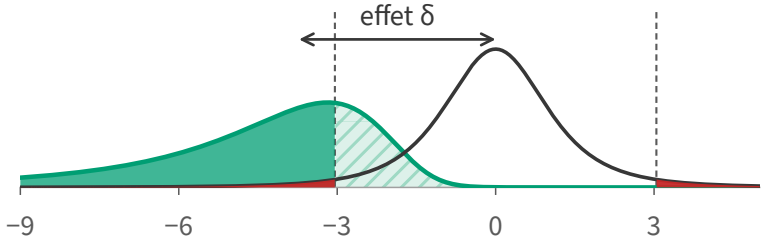
Annexe A5 – Puissance, effet et nombre d'éprouvettes

Une image par clic : $n = 3$ à 105 à effet fixe, puis effets plus grands à $n = 3$.

retour à la diapo 21

n = 3 éprouvettes par échantillon **puissance 1 − β = 73,0 %**

taille d'effet δ/σ = 3,0

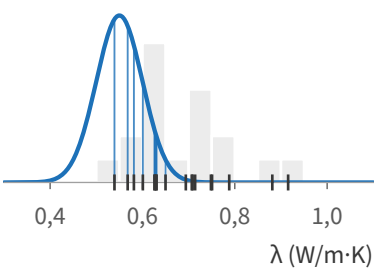


Annexe A6 – Le maximum de vraisemblance

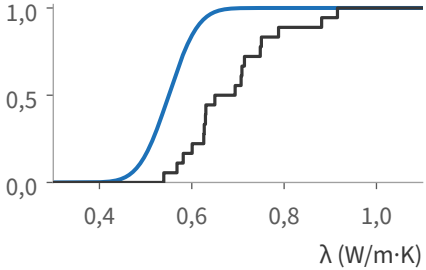
Une image par clic : la loi normale s'ajuste, ln L monte.

retour à la diapo 25

densité de la loi normale essayée



fonction de répartition



vraisemblance



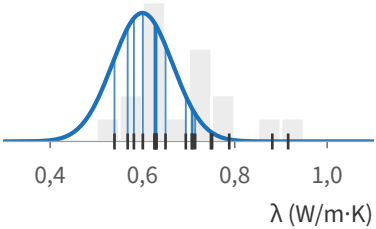
essais successifs

Annexe A6 – Le maximum de vraisemblance

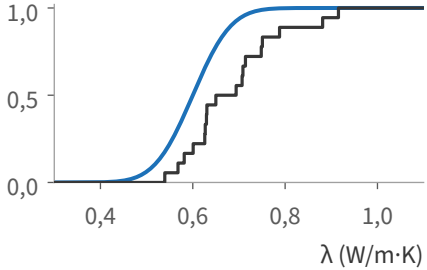
Une image par clic : la loi normale s'ajuste, ln L monte.

retour à la diapo 25

densité de la loi normale essayée



fonction de répartition



vraisemblance

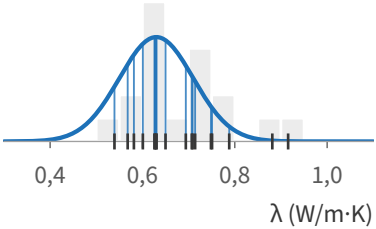


Annexe A6 – Le maximum de vraisemblance

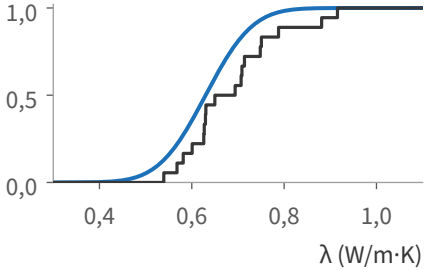
Une image par clic : la loi normale s'ajuste, ln L monte.

retour à la diapo 25

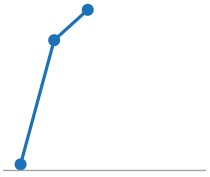
densité de la loi normale essayée



fonction de répartition



vraisemblance



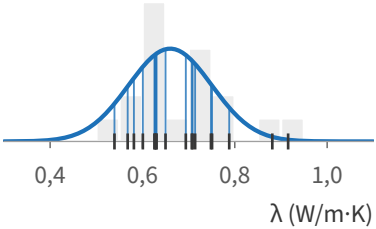
essais successifs

Annexe A6 – Le maximum de vraisemblance

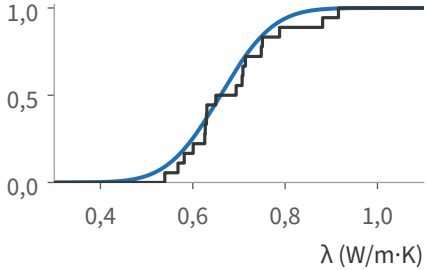
Une image par clic : la loi normale s'ajuste, ln L monte.

retour à la diapo 25

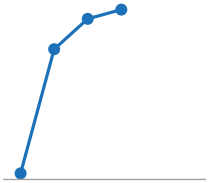
densité de la loi normale essayée



fonction de répartition



vraisemblance



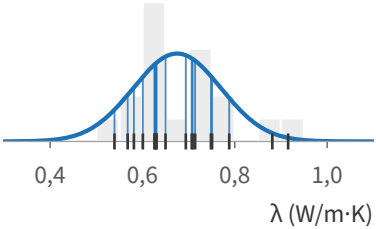
essais successifs

Annexe A6 – Le maximum de vraisemblance

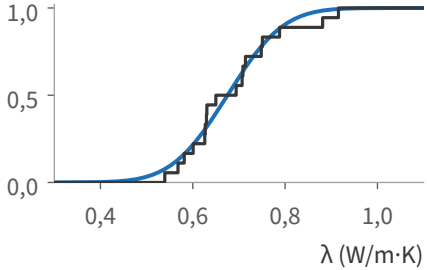
Une image par clic : la loi normale s'ajuste, ln L monte.

retour à la diapo 25

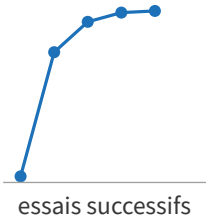
densité de la loi normale essayée



fonction de répartition



vraisemblance

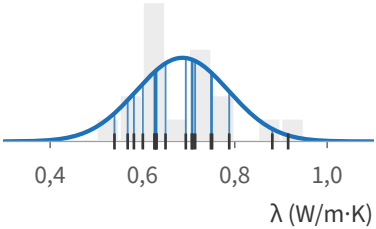


Annexe A6 – Le maximum de vraisemblance

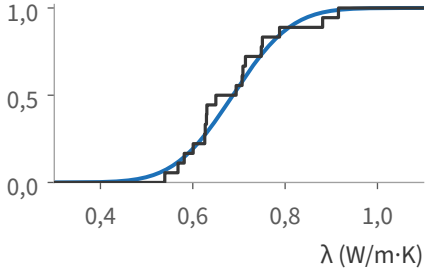
Une image par clic : la loi normale s'ajuste, ln L monte.

retour à la diapo 25

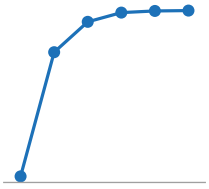
densité de la loi normale essayée



fonction de répartition



vraisemblance



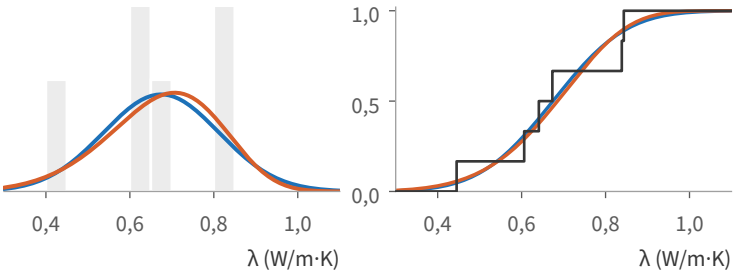
essais successifs

Annexe A7 – L'AICc selon le nombre d'éprouvettes

Une image par clic : n = 6, 12, 18, 30, 50, 80, 132.

retour à la diapo 27

n = 6 éprouvettes



	normale	Weibull
KS : p	0,89	0,86
AICc	1,2	1,0

verdict de l'AICc :

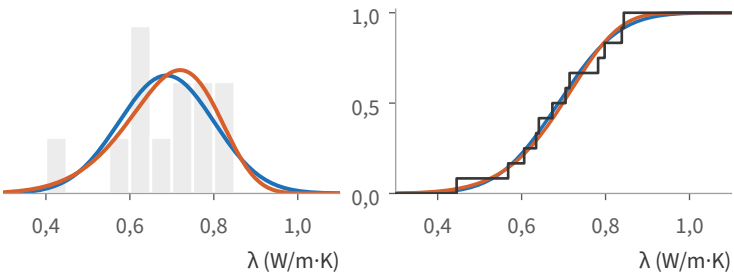
pas de préférence (écart < 2)

Annexe A7 – L'AICc selon le nombre d'éprouvettes

Une image par clic : n = 6, 12, 18, 30, 50, 80, 132.

retour à la diapo 27

n = 12 éprouvettes



	normale	Weibull
KS : p	0,97	0,98
AICc	−12,9	−13,5

verdict de l'AICc :

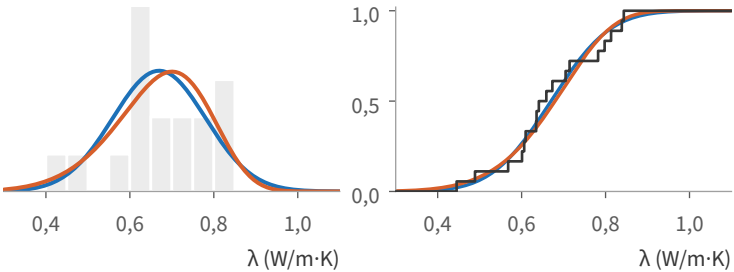
pas de préférence (écart < 2)

Annexe A7 – L'AICc selon le nombre d'éprouvettes

Une image par clic : n = 6, 12, 18, 30, 50, 80, 132.

retour à la diapo 27

n = 18 éprouvettes



	normale	Weibull
KS : p	0,92	0,87
AICc	−23,3	−23,3

verdict de l'AICc :

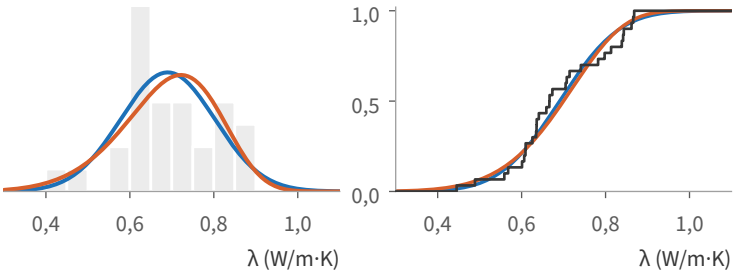
pas de préférence (écart < 2)

Annexe A7 – L'AICc selon le nombre d'éprouvettes

Une image par clic : n = 6, 12, 18, 30, 50, 80, 132.

retour à la diapo 27

n = 30 éprouvettes



	normale	Weibull
KS : p	0,68	0,42
AICc	−41,6	−41,0

verdict de l'AICc :

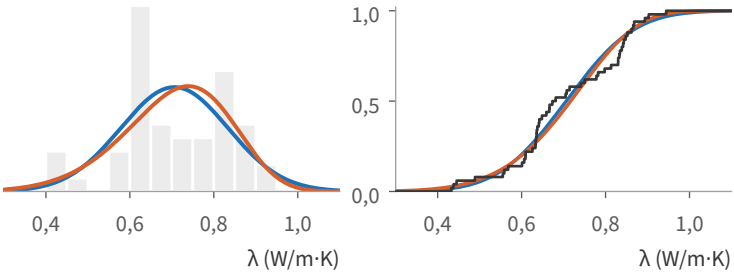
pas de préférence (écart < 2)

Annexe A7 – L'AICc selon le nombre d'éprouvettes

Une image par clic : n = 6, 12, 18, 30, 50, 80, 132.

retour à la diapo 27

n = 50 éprouvettes



	normale	Weibull
KS : p	0,31	0,34
AICc	−59,3	−60,3

verdict de l'AICc :

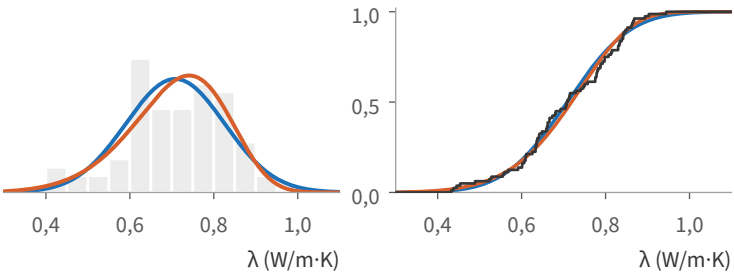
pas de préférence (écart < 2)

Annexe A7 – L'AICc selon le nombre d'éprouvettes

Une image par clic : n = 6, 12, 18, 30, 50, 80, 132.

retour à la diapo 27

n = 80 éprouvettes



	normale	Weibull
KS : p	0,34	0,63
AICc	−111,0	−114,2

verdict de l'AICc :

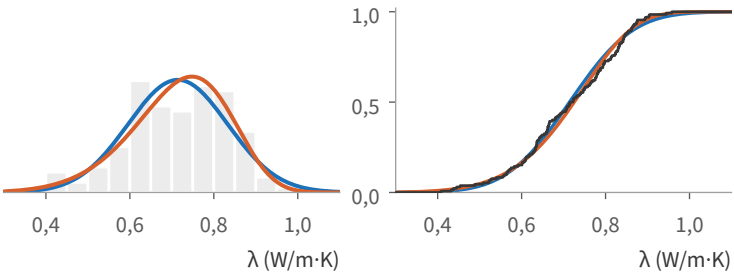
Weibull préférée

Annexe A7 – L'AICc selon le nombre d'éprouvettes

Une image par clic : n = 6, 12, 18, 30, 50, 80, 132.

retour à la diapo 27

n = 132 éprouvettes



	normale	Weibull
KS : p	0,42	0,38
AICc	−183,9	−188,9

verdict de l'AICc :

Weibull préférée