



Éléments finis et Monte Carlo

 Jérémy Aubé-Lévèsque

 29 août 2024

 GIREF, Université Laval

➤ Méthode des éléments finis

➤ Monte Carlo

➤ Bootstrap

➤ Application

Méthode des éléments finis

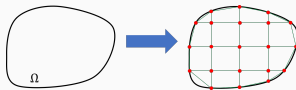
Permet de résoudre numériquement des EDOs et EDPs avec ou sans solution analytique.

Exemple

$$\Omega = [0, 1] \times [0, 1]$$

$$\begin{cases} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f & \text{dans } \Omega \\ u = g & \text{sur } \partial\Omega \end{cases}$$

Discrétisation (maillage)



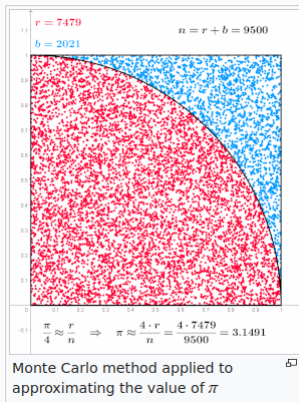
- Chaque sommet est un degré de liberté.
- La solution approximative $U = [u_1, \dots, u_n]^T$, u_i est la valeur approximative de u au sommet i .

Monte Carlo

La méthode de Monte Carlo cherche à estimer les statistiques d'une v.a. d'espérance finie par simulation. La méthode est une application de la loi forte des grands nombres.

Etape simulation de monte carlo

1. Définir un domaine de valeurs d'entrée possibles.
2. Générer aléatoirement des valeurs d'entrée à partir d'une distribution de probabilité sur ce domaine.
3. Effectuer un calcul déterministe des résultats.
4. Analyse statistique des résultats.



Rappel

On dit que (X_n) converge presque sûrement vers X et on note $X_n \xrightarrow{\text{a.s.}} X$ si

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} X_n = X\right) = 1$$

On dit que (X_n) converge en probabilité vers X et on note $X_n \xrightarrow{p} X$ si

$$\forall \varepsilon > 0, \quad \lim_{n \rightarrow \infty} \mathbb{P}(|X_n - X| \geq \varepsilon) = 0.$$

Loi forte des grands nombres Si $(X_n)_{n \geq 0}$ est une suite de variables aléatoires indépendantes identiquement distribuées, on a équivalence entre :

1. $\mathbb{E}(|X_1|) < +\infty$,
2. La suite $\frac{X_1 + \dots + X_n}{n}$ converge presque sûrement.

Si l'une de ces deux conditions est vraie alors, $\frac{X_1 + \dots + X_n}{n} \xrightarrow{\text{a.s.}} \mathbb{E}[X_1]$

Exemple

Soit $X \sim \gamma = \mathcal{N}(0, 1)$ et $a \in \mathbb{R}$. On définit :

$$Y := \frac{1}{a - X}$$

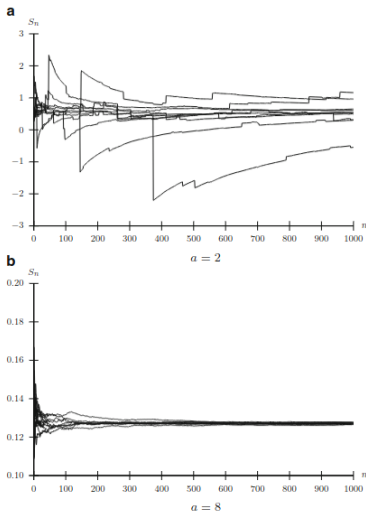
Y est fini, mais $\mathbb{E}_\gamma(Y)$ n'est pas fini. Ainsi, on a :

$$\int_{\mathbb{R}} |Y| \, d\gamma = \int_{\mathbb{R}} \left| \frac{1}{a - x} \right| \gamma(x) \, dx \geq \int_A \left| \frac{1}{a - x} \right| \gamma(x) \, dx \geq \inf_{x \in A} \gamma(x) \cdot \int_{\mathbb{R}} \left| \frac{1}{a - x} \right| \, dx$$

On sait que $Y \notin L^1$

► Monte Carlo ► Hypothèse d'espérance finie

Simulation monte carlo $\mathbb{E}_\gamma(Y)$:



► Monte Carlo ► application aux éléments finis



Soit $\theta^{(1)}, \theta^{(2)} \sim F_1, F_2$ et deux échantillons $\theta_1^{(1)}, \dots, \theta_n^{(1)} \stackrel{\text{i.i.d.}}{\sim} F_1$, $\theta_1^{(2)}, \dots, \theta_n^{(2)} \stackrel{\text{i.i.d.}}{\sim} F_2$

$$\begin{cases} \theta^{(1)} \frac{\partial^2 u}{\partial x^2} + \theta^{(2)} \frac{\partial^2 u}{\partial y^2} = f & \text{dans } \Omega \\ u = g & \text{sur } \partial\Omega \end{cases}$$

La solution U devient donc, une solution aléatoire $U(x, y, \theta^{(1)}, \theta^{(2)})$ et on peut estimer ses statistiques.

$$\mathbb{E}^{MC}[U] = \frac{1}{N} \sum_{i=1}^N U(\theta_i^{(1)}, \theta_i^{(2)})$$

$$\text{Var}^{MC}[U] = \frac{1}{N-1} \sum_{i=1}^N \left(U(\theta_i^{(1)}, \theta_i^{(2)}) - \mathbb{E}^{MC}[U] \right)^2$$

$$\text{COV}^{MC} = \frac{\sqrt{\text{Var}^{MC}[U]}}{\sqrt{N} \mathbb{E}^{MC}[U]}$$

Pour une simulation efficace, il faut que $COV \leq 0.05$. Le COV (Coefficient of variation) est le ratio entre l'écart type de l'estimateur de $\mathbb{E}^{MC}[U]$ et sa valeur. Quand $COV \leq 0.05$, augmenter le nombre de simulations ne change plus significativement la valeur de l'estimateur.

Dans les simulations de Monte Carlo, les statistiques calculées sont des estimateurs, car elles proviennent d'un échantillon fini, étant donné que la distribution de la solution aléatoire est inconnue. Cela introduit de l'incertitude dans nos résultats.

Definition

Un intervalle de confiance est un intervalle censé contenir un paramètre inconnu que l'on cherche à estimer avec un certain niveau de confiance.

Soit $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} F$ et un IC $I_\alpha = [T_1, T_2] \subseteq \mathbb{R}$ de niveau de confiance $(1 - 2\alpha) \times 100\%$ pour le paramètre $\theta = \theta(F)$. Alors,

$$P(\theta \in I_\alpha) = 1 - 2\alpha.$$

Exemple d'Intervalle de Confiance Paramétrique

Soit $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$ et $\hat{\theta} = \bar{X}$, alors il est connu que

$$\frac{\bar{x} - \mu}{\hat{S}E} \sim t_{n-1}.$$

Il est donc facile de construire un intervalle de confiance pour μ :

$$\left[\bar{x} - t_{n-1}^{-1}(1 - \alpha) \hat{S}E, \bar{x} - t_{n-1}^{-1}(\alpha) \hat{S}E \right]$$

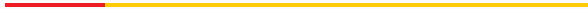
Approche Asymptotique

si l'on sait $\frac{\hat{\theta} - \theta}{\hat{S}E} \sim G_n$ et qu'on sait que $G_n \rightarrow F$ lorsque $n \rightarrow \infty$ et que l'on connaît F . Alors, on peut approximer l'intervalle de confiance par

$$P \left(\theta \in \left[\hat{\theta} - F^{-1}(1 - \alpha) \hat{S}E, \hat{\theta} - F^{-1}(\alpha) \hat{S}E \right] \right) \approx 1 - 2\alpha.$$

Dans le monde réel, les hypothèses paramétriques ou asymptotiques ne sont pas toujours vérifiées. Par conséquent, nous utilisons des intervalles de confiance non paramétriques, comme ceux basés sur la méthode du bootstrap

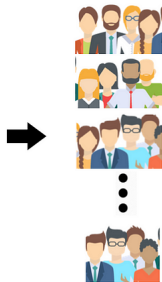
Bootstrap



Sample Size N



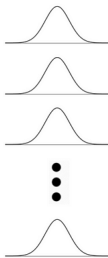
B Bootstrap
Samples, each size n



Estimate θ

θ
 θ
 θ
 \vdots
 θ

Inference



Definition

La fonction indicatrice d'un sous-ensemble A d'un ensemble X est une fonction

$$\mathbf{1}_A : X \rightarrow \{0, 1\}$$

définie par

$$\mathbf{1}_A(x) := \begin{cases} 1 & \text{si } x \in A, \\ 0 & \text{si } x \notin A. \end{cases}$$

Definition

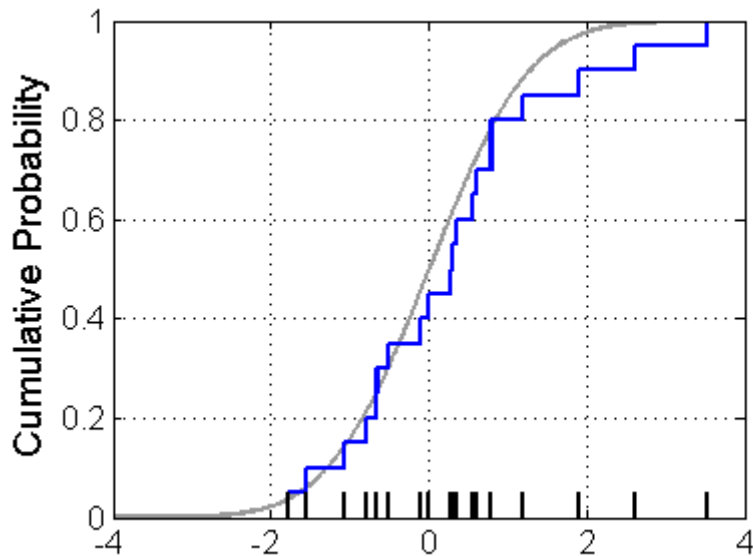
Soit $X_1, \dots, X_n \stackrel{\text{i.i.d.}}{\sim} F$ représentant un échantillon de F . La fonction de répartition empirique $\hat{F}_n(x)$ est définie par :

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{X_i \leq x\}}, \forall x \in \mathbb{R}$$

- Loi de probabilité discrète en associant à chaque X_i , la masse $\frac{1}{n}$
- $\forall x \in \mathbb{R}, \mathbf{1}_{\{X_i \leq x\}} \sim \text{Bernoulli}(F(x))$
- $\hat{F}_n(x) \xrightarrow{\text{a.s.}} F(x)$ d'après la loi forte des grands nombres

Pour les visuels :

$$\mathbb{P} \left(\lim_{n \rightarrow \infty} \hat{F}_n(x) = \lim_{n \rightarrow \infty} \frac{\mathbf{1}_{\{X_1 \leq x\}} + \dots + \mathbf{1}_{\{X_n \leq x\}}}{n} = \mathbb{E} [\mathbf{1}_{\{X_1 \leq x\}}] = F(x) \right) = 1$$



Principe du plug-in : $\theta(\hat{F}_n) \approx \theta(F)$ lorsque $\hat{F}_n \approx F$, pourvu que θ soit continue par rapport à la convergence a.s.

Exemple La moyenne d'une variable $X \sim F$ (quand elle existe) peut s'écrire :

$$\mathbb{E}_F(X) = \theta(F) = \int x dF(x).$$

En appliquant le principe du plug-in :

$$\mathbb{E}_{\hat{F}_n}(X) = \theta(\hat{F}_n) = \int x d\hat{F}_n(x) = \sum_{i=1}^n X_i \cdot \frac{1}{n} = \bar{X},$$

La variance, si elle existe, est :

$$\text{Var}_F(X) = \theta(F) = \int \left[x - \int y dF(y) \right]^2 dF(x).$$

Plug-in :

$$\text{Var}_{\hat{F}_n}(X) = \theta(\hat{F}_n) = \int \left[x - \int y d\hat{F}_n(y) \right]^2 d\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2.$$

On va nommer la variance d'un estimateur obtenue par le plug-in $v_{\text{Boot}} := \text{Var}_{\hat{F}_n}(\hat{\theta})$, l'estimateur bootstrap idéal. Nous verrons plus tard qu'il n'est pas toujours possible de calculer cette valeur et nous l'approximerons par Monte Carlo.

Exemple Soit $X_1, X_2, \dots, X_n \stackrel{\text{i.i.d.}}{\sim}$ et $\hat{\theta} = \bar{X}_n$ telle que $\text{Var}(X_1) = \sigma^2$. Alors on sait que

$$\text{Var}_F(\bar{X}_n) = \text{Var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{\sigma^2}{n} = \frac{1}{n} \int \left[x - \int y dF(y) \right]^2 dF(x).$$

Plug-in :

$$\text{Var}_F(\bar{X}_n) \approx \text{Var}_{\hat{F}_n}(\bar{X}_n) = \frac{1}{n} \int \left[x - \int y d\hat{F}(y) \right]^2 d\hat{F}(x) = \frac{1}{n^2} \sum_{i=1}^n (X_i - \bar{X})^2.$$

et donc, on a que $v_{\text{Boot}} = \frac{1}{n^2} \sum_{i=1}^n (X_i - \bar{X})^2$

Definition

un échantillon bootstrap sont les réalisations de $X_1^*, X_2^*, \dots, X_n^* \stackrel{\text{i.i.d.}}{\sim} \hat{F}_n$. Par définition,

$$P(X_i^* = x_j) = \frac{1}{n}, \quad 1 \leq i, j \leq n$$

Pour certaines statistique, il n'existe aucune formule analytique pour trouver v_{Boot} comme à la diapositive précédente. Par contre, on sait que v_{Boot} à toujours la forme d'une somme finie puisque l'on sait que $X_1^*, X_2^*, \dots, X_n^*$ suivent une loi conditionnelle discrète à $(X_1, \dots, X_n) = (x_1, \dots, x_n)$. On peut donc, calculer v_{Boot} en faisant varier $(X_1^*, X_2^*, \dots, X_n^*)$ dans l'ensemble fini de tous les échantillons bootstrap possibles.

Alors, on peut écrire v_{Boot} :

$$\begin{aligned}
 v_{\text{Boot}} &= \text{Var}_{\hat{F}_n}(\hat{\theta}) \\
 &= \mathbb{E}_{\hat{F}_n} \left[\hat{\theta}(X_1^*, \dots, X_n^*) - \mathbb{E}_{\hat{F}_n} \left(\hat{\theta}(X_1^*, \dots, X_n^*) \right) \right]^2 \\
 &= \sum_{i=1}^N \left[\hat{\theta}_i(X_1^*, \dots, X_n^*) - \sum_{j=1}^N \hat{\theta}_j(X_1^*, \dots, X_n^*) w_j \right]^2 w_j \\
 &= \text{Var}(\hat{\theta}(X_1^*, \dots, X_n^*) \mid X_1, \dots, X_n)
 \end{aligned}$$

où $\hat{\theta}(X_1^*, \dots, X_n^*)$ représente les valeurs de l'estimateur pour un certain échantillon bootstrap ,
 w_j sa probabilité et N le nombre total de réalisations bootstrap distincts.

Il y a $m = \binom{2n-1}{n}$ échantillons bootstrap possibles si :

- les x_i sont distincts,
- on ne prend pas en compte les permutations.

soit $(X_1, \dots, X_n) = (x_1, \dots, x_n)$ et $A_i = \#\{X_j^* = x_i, 1 \leq j \leq n\}$ pour $i = 1, \dots, n$ on peut alors définir un échantillon bootstrap $(A_1, \dots, A_n) = (a_1, \dots, a_n)$ tel que

$$\forall a_i \geq 0, \text{ et } a_1 + \dots + a_n = n$$

alors, la probabilité de sélectionner un échantillon bootstrap suit la probabilité multinomial

$$\Pr(A_1 = a_1, \dots, A_n = a_n \mid X_1, \dots, X_n) = \frac{n!}{a_1! \dots a_n!} \cdot \left(\frac{1}{n}\right)^{a_1} \dots \left(\frac{1}{n}\right)^{a_n} = \frac{n!}{n^n a_1! \dots a_n!}$$

ce qui est équivalent

$$\Pr(\hat{\theta}^* = \hat{\theta}^*(a_1, \dots, a_n) \mid X_1, \dots, X_n) = \frac{n!}{n^n a_1! \dots a_n!}$$

La plus grande probabilité :

- Tous les $a_i = 1$,

$$Pr(\hat{\theta}^* = \hat{\theta} \mid X_1, \dots, X_n) = \frac{n!}{n^n}$$
- Comme on peut voir, m devient très grand. Dans la pratique, on ne calculera v_{Boot} de façon exacte, mais plutôt par Monte Carlo.
- la probabilité de rééchantillonner le même échantillon est faible. Donc, comme si on rééchantillonnait directement dans F

n	m	$Pr(\hat{\theta}^* = \hat{\theta} \mid X_1, \dots, X_n)$
2	3	0.5
3	10	0.2222
4	35	0.0940
5	126	0.0384
6	462	1.5×10^{-2}
10	9.2×10^4	3.6×10^{-4}
15	7.8×10^7	3×10^{-6}
20	6.9×10^{10}	2.3×10^{-8}

Approximation de v_{Boot} par Monte Carlo :

1) Tirer B échantillons bootstrap $X_{1b}^*, X_{2b}^*, \dots, X_{nb}^* \stackrel{\text{i.i.d.}}{\sim} \hat{F}_n$ pour $b = 1, \dots, B$ puis calculer $\hat{\theta}_b^* = \hat{\theta}(X_{1b}^*, \dots, X_{nb}^*)$.

2) calculez :

$$v_{\text{Boot}}^{(B)} = \frac{1}{B} \sum_{b=1}^B \left(\hat{\theta}_b^* - \hat{\theta}_{(\cdot)}^* \right)^2 = \frac{1}{B} \sum_{b=1}^B \left(\hat{\theta}_b^* \right)^2 - \left(\hat{\theta}_{(\cdot)}^* \right)^2,$$

$$\text{où } \hat{\theta}_{(\cdot)}^* = \frac{1}{B} \sum_{b=1}^B \hat{\theta}_b^*$$

On sait que $v_{\text{Boot}}^{(B)} \xrightarrow{\text{a.s.}} v_{\text{Boot}}$ par la loi forte des grands nombres. Donc, $v_{\text{Boot}}^{(B)} \approx v_{\text{Boot}}$, dès que B est assez grand.

$G_n = \Pr(R(X_1, \dots, X_n, F) \leq x)$ est la distribution de la v.a R et

$G_{Boot} = \Pr(R(X_1^*, \dots, X_n^*, \hat{F}_n) \leq x)$ est la distribution bootstrap de la v.a R . Par exemple, $R(X_1, \dots, X_n, F) = \hat{\theta}(X_1, \dots, X_n)$ ou $R(X_1, \dots, X_n, F) = \hat{\theta}(X_1, \dots, X_n) - \mu$.

Définition Soit ρ une métrique sur $\mathcal{F}_{\mathbb{R}^s} = \{\text{toutes les distributions de } \mathbb{R}^s\}$. G_{Boot} est faiblement ρ -consistant si $\rho(G_{Boot}, G_n) \xrightarrow{P} 0$ lorsque $n \rightarrow \infty$. G_{Boot} est fortement ρ -consistant si $\rho(G_{Boot}, G_n) \xrightarrow{\text{a.s.}} 0$ lorsque $n \rightarrow \infty$.

Distance de Mallows :

$$\mathcal{F}_{\mathbb{R}^s} = \{G \in \mathcal{F}_{\mathbb{R}^s} : \int \|X\|^r dG(x) < +\infty\}$$

$$\tilde{\rho}_r(H, G) = \inf_{\mathcal{T}_{X,Y}} (\mathbb{E}[\|x - y\|^r])^{1/r}$$

$\mathcal{T}_{X,Y}$ représente l'ensemble des distributions jointes de (X, Y) tel que leur distribution marginal sont H et G

Propriétés présentées par Bickel et Freedman dans leur article publié en 1981 *Some Asymptotic Theory for the Bootstrap*

$$\tilde{\rho}_r(\hat{F}_n, F) \xrightarrow{\text{a.s.}} 0, F \in \mathcal{F}_{\mathbb{R}^s} \quad (1)$$

soit U et V des vecteurs aléatoires et $a \in \mathbb{R}$

$$\tilde{\rho}_r(aU, aV) = |a| \tilde{\rho}_r(U, V) \quad (2)$$

si $\mathbb{E}[\|U\|^2] < +\infty$ et $\mathbb{E}[\|V\|^2] < +\infty$ alors,

$$\tilde{\rho}_2(U, V)^2 = \tilde{\rho}_2(U - \mathbb{E}[U], V - \mathbb{E}[V])^2 + \|\mathbb{E}[U] - \mathbb{E}[V]\|^2 \quad (3)$$

soit $\{U_j\}, \{V_j\} \subset \mathcal{F}_{\mathbb{R}^s}$ deux suites indépendantes de vecteurs aléatoires et $\mathbb{E}[U_j] = \mathbb{E}[V_j], \forall j$

$$\tilde{\rho}_2 \left(\sum_{j=1}^m U_j, \sum_{j=1}^m V_j \right)^2 \leq \sum_{j=1}^m \tilde{\rho}_2(U_j, V_j)^2 \quad (4)$$

Exemple : soit $F \in \mathcal{F}_{\mathbb{R}^s}$ on pose $R = \sqrt{n}(\bar{X}_n - \mu)$ et $R^* = \sqrt{n}(\bar{X}_n^* - \bar{X}_n)$

$$\begin{aligned} \tilde{\rho}_2(G_{Boot}, G_n) &= \tilde{\rho}_2(\sqrt{n}(\bar{X}_n^* - \bar{X}_n), \sqrt{n}(\bar{X}_n - \mu)) \\ &= \frac{1}{\sqrt{n}} \tilde{\rho}_2\left(\sum_{j=1}^n (X_j^* - \bar{X}_n), \sum_{j=1}^n (X_j - \mu)\right), \text{ par (2)} \\ &\leq \sqrt{\frac{1}{n} \sum_{j=1}^n (\tilde{\rho}_2(X_j^* - \bar{X}_n, X_j - \mu))^2}, \text{ par (4)} \\ &= \tilde{\rho}_2(X_1^* - \bar{X}_n, X_1 - \mu) \\ &= \sqrt{\tilde{\rho}_2(X_1^*, X_1)^2 - \|\mu - \bar{X}_n\|^2}, \text{ par (3)} \\ &= \sqrt{\tilde{\rho}_2(\hat{F}_n, F)^2 - \|\mu - \bar{X}_n\|^2} \\ &\rightarrow 0, \text{ par (1)} \end{aligned}$$

On vient de montrer la forte $\tilde{\rho}_2$ -consistance de G_{Boot} pour $\hat{\theta} = \bar{X}_n$

Définition

Un pivot est une v.a $R(X_1, \dots, X_n, F) \sim G_n$ indépendante de F

Exemple

soit $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$ alors,

$$R(X_1, \dots, X_n, F) = \frac{\bar{x} - \mu}{\hat{s}e} \sim t_{n-1}.$$

Peut importer μ et σ^2

Définition

Soit $X \sim F$ et $\theta(F)$ est un paramètre de localisation si :

- i) $\theta(X + b) = \theta(X) + b$, quelle que soit b constante ;
- ii) $X \geq 0$ implique que $\theta(X) \geq 0$;
- iii) $\theta(aX) = a\theta(X)$, quelle que soit $a > 0$ constante.

Par exemple, la moyenne, la médiane, les quartiles et etc sont des paramètres de localisation

Si θ est un paramètre de localisation, alors la v.a $\frac{\hat{\theta}_n - \theta}{\hat{SE}} \sim G_n$ est souvent un pivot exact ou approximatif.

$$\Pr \left(G_n^{-1}(\alpha) \leq \frac{\hat{\theta}_n - \theta}{\hat{SE}} \leq G_n^{-1}(1 - \alpha) \right) = 1 - 2\alpha$$

$$\Pr \left(\hat{\theta}_n - G_n^{-1}(1 - \alpha)\hat{SE} \leq \theta \leq \hat{\theta}_n - G_n^{-1}(\alpha)\hat{SE} \right) = 1 - 2\alpha$$

Ce qui donne comme I.C $\left[\hat{\theta}_n - G_n^{-1}(1 - \alpha)\hat{SE}, \theta - \hat{\theta}_n - G_n^{-1}(\alpha)\hat{SE} \right]$ à $(1 - 2\alpha) \times 100\%$

Bootstrap-t

On approxime $G_n \approx \hat{t}_{Boot}(x) = \Pr(R(X_1^*, \dots, X_n^*, \hat{F}_n) \leq x)$ et $R(X_1^*, \dots, X_n^*, \hat{F}_n) = \frac{\hat{\theta}_n^* - \hat{\theta}_n}{\hat{SE}^*}$
et l'intervalle de confiance calculer par le bootstrap-t est

$$\left[\hat{\theta}_n - \hat{t}_{Boot}^{-1}(1 - \alpha) \cdot \hat{SE}, \hat{\theta}_n - \hat{t}_{Boot}^{-1}(\alpha) \cdot \hat{SE} \right]$$

Algorithme

1. tirer B échantillons bootstrap $X_{1b}^*, X_{2b}^*, \dots, X_{nb}^* \stackrel{\text{i.i.d.}}{\sim} \hat{F}_n$ pour $b = 1, \dots, B$, et calculer $\hat{\theta}_b^* = \hat{\theta}(X_{1b}^*, \dots, X_{nb}^*)$

2. de chaque échantillon $X_{1b}^*, X_{2b}^*, \dots, X_{nb}^* \stackrel{\text{i.i.d.}}{\sim} \hat{F}_n$ tirer B_2 échantillons bootstrap $X_{1bj}^*, X_{2bj}^*, \dots, X_{nbj}^* \stackrel{\text{i.i.d.}}{\sim} \hat{F}_{nb}$ pour $j = 1, \dots, B_2$ pour calculer $\hat{\theta}_{bj}^*$, $j = 1, \dots, B_2$, $\bar{\theta}_b^* = \frac{1}{B_2} \sum_{j=1}^{B_2} \hat{\theta}_{bj}^*$, et calculer

$$\hat{\sigma}_b^{*(B_2)} = \sqrt{\frac{1}{B_2} \sum_{j=1}^{B_2} (\hat{\theta}_{bj}^* - \bar{\theta}_b^*)^2}.$$

Il est préférable de calculer $\hat{\sigma}_b$ analytiquement avec le plug-in lorsque c'est possible sinon le nombre de calcul est très élevé ($B \cdot B_2$).

Definition

Supposons $\hat{\theta}$ être un estimateur de θ et $\hat{\theta}^*$ l'estimateur bootstrap basé sur $X_1^*, X_2^*, \dots, X_n^* \stackrel{\text{i.i.d.}}{\sim} \hat{F}_n$.

L'IC par la méthode des percentiles a un niveau de confiance de $1 - 2\alpha$ est

$$[G_{\text{Boot}}^{-1}(\alpha), G_{\text{Boot}}^{-1}(1 - \alpha)]$$

Démonstration

supposons qu'il existe φ une transformation inversible et croissante tel que

$$Pr(\varphi(\hat{\theta}) - \varphi(\theta) \leq x) = \Psi(x)$$

tel que $\Psi(x)$ soit symétrique, continue, indépendante de F et donc, un pivot

On pose $z_\alpha = \Psi^{-1}(\alpha)$ et $z_{1-\alpha} = \Psi^{-1}(1 - \alpha) = -z_\alpha$ (par symetrie)

alors, nous avons

$$\Pr \left(z_\alpha \leq \varphi(\hat{\theta}) - \varphi(\theta) \leq -z_\alpha \right) = 1 - 2\alpha$$

$$\Pr \left(z_\alpha + \varphi(\hat{\theta}) \leq \varphi(\theta) \leq -z_\alpha + \varphi(\hat{\theta}) \right) = 1 - 2\alpha$$

$$\Pr \left(\varphi^{-1}(z_\alpha + \hat{\theta}) \leq \theta \leq \varphi^{-1}(-z_\alpha + \hat{\theta}) \right) = 1 - 2\alpha$$

Puisque l'on sait que $\varphi(\hat{\theta}^*) - \varphi(\hat{\theta}) \approx \varphi(\hat{\theta}) - \varphi(\theta)$

alors, on a que

$$\Pr \left(\varphi^{-1}(z_\alpha + \hat{\theta}) \leq \hat{\theta}^* \leq \varphi^{-1}(-z_\alpha + \hat{\theta}) \right) \approx 1 - 2\alpha$$

On sait que $\Pr \left(G_{\text{Boot}}^{-1}(\alpha) \leq \hat{\theta}^* \leq G_{\text{Boot}}^{-1}(1 - \alpha) \right) = 1 - 2\alpha$

Donc, on vient de demontrer que $G_{\text{Boot}}^{-1}(\alpha) \approx \varphi^{-1}(z_\alpha + \hat{\theta})$ et $G_{\text{Boot}}^{-1}(1 - \alpha) \approx \varphi^{-1}(-z_\alpha + \hat{\theta})$

Lorsque l'hypothèse de transformation φ est approximativement vérifiée pour un grand n , l'intervalle de confiance obtenu par percentiles est asymptotiquement valide, et sa performance dépendra de la qualité de cette approximation.

Pour relacher cette hypothese, il y a la methode BC_a qui fait l'hypothese :

$$\exists a, z_0 \in \mathbb{R} \text{ tel que } P\left(\frac{\varphi(\hat{\theta}) - \varphi(\theta)}{1 + a\varphi(\theta)} + z_0 \leq x\right) = \Phi(x)$$

Comparaison des méthodes

Bootstrap-t : la distribution de $\hat{\theta}_n$ doit être symétrique par rapport à θ . Si ce n'est pas le cas, on doit trouver une transformation inversible et croissante pour transformer nos données et ensuite appliquer la méthode.

Percentile : ne performe pas bien lorsque la distribution de la population est asymétrique et lorsque n est faible. Il n'est pas nécessaire de trouver une transformation.

Les tableaux montrent le nombre de fois où la statistique θ (qui est connue) ne se trouve pas dans les intervalles de confiance calculés selon différentes méthodes. Nous avons effectué 10000 simulations avec les paramètres suivants :

- $X \sim \mathcal{N}(0, 1)$
- $\hat{\theta} = \exp(\bar{X})$
- $n = 10$
- Intervalles de confiance à 95%

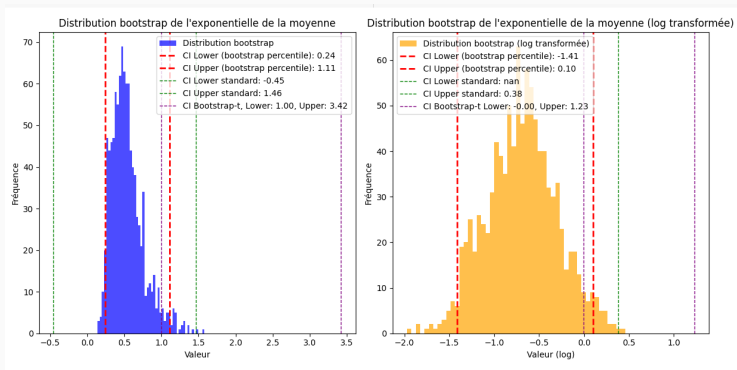
Les résultats sont présentés dans le tableau ci-dessous :

Le temps d'exécution du code est de 0.299335 heures

	% miss left	% miss right	% total
t-Student	5.65	0.60	6.25
percentile (non paramétric)	5.03	4.93	9.96
bootstrap-t (non paramétrique)	97.80	0.00	97.80



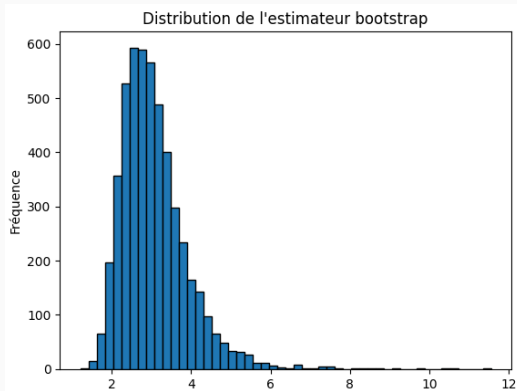
Aperçu de la simulation présentée dans la diapositive précédente. Les graphiques montrent la distribution approximative de $\hat{\theta}^*$ avec et sans transformation logarithmique, ainsi que les intervalles de confiance calculés selon différentes méthodes. Ces graphiques illustrent simplement les résultats d'une simulation.



Les percentiles (lignes rouges) sont clairement plus efficaces puisque la population est symétrique et l'asymétrie de la distribution de $\hat{\theta}$ n'affecte pas la performance de la méthode.

Nous avons effectué 5000 simulations avec les paramètres suivants :

- $X \sim \text{LogNormal}(0, \frac{3}{2})$
- $\hat{\theta} = \bar{X}$
- $n = 100$
- Intervalles de confiance à 95%



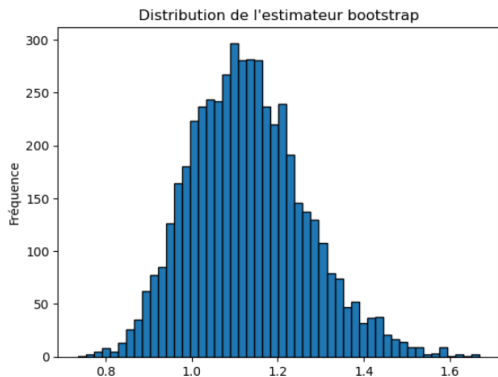
Le temps d'exécution du code est de 0.647093 heures

	% miss left	% miss right	% total
percentile (non paramétrique)	2.28	9.30	11.58
bootstrap-t (non paramétrique)	1.12	7.22	8.34

Étant donné que la distribution de X est fortement asymétrique et que la distribution approximative de $\hat{\theta}^*$ présente également une forte asymétrie, les deux IC non paramétriques ne donnent pas de bons résultats.

Nous avons effectué 5000 simulations avec les paramètres suivants :

- $X \sim \text{LogNormal}(0, \frac{1}{2})$
- $\hat{\theta} = \bar{X}$
- $n = 20$
- Intervalles de confiance à 95%



Le temps d'exécution du code est de 0.472982 heures

	% miss left	% miss right	% total
student (paramétrique)	0.66	6.32	6.98
percentile (non paramétrique)	3.12	5.92	9.04
bootstrap-t (non paramétrique)	1.66	4.00	5.66

Étant donné que la distribution de X est asymétrique et que la distribution approximative de $\hat{\theta}^*$ présente une légère asymétrie, les deux IC non paramétriques ne donnent pas de bons résultats. Mais, le Bootstrap-t est le mieux dans cette situation.

Application

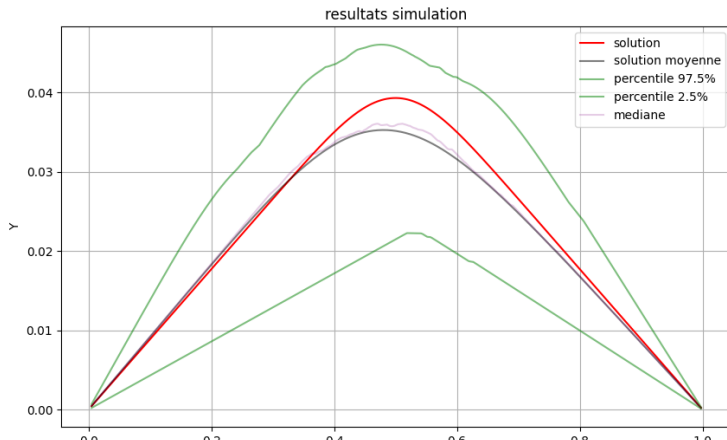
Premiere application de difusion en 1D

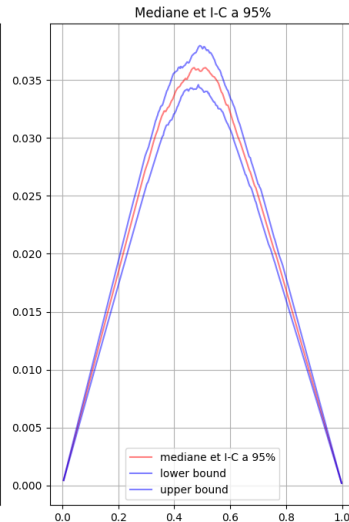
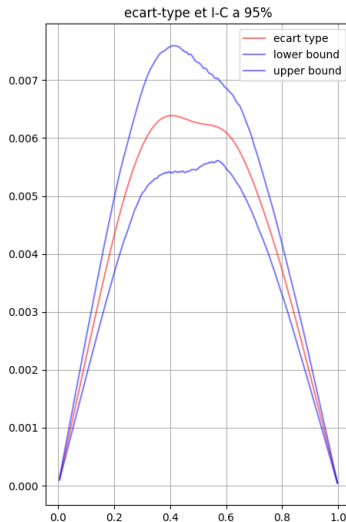
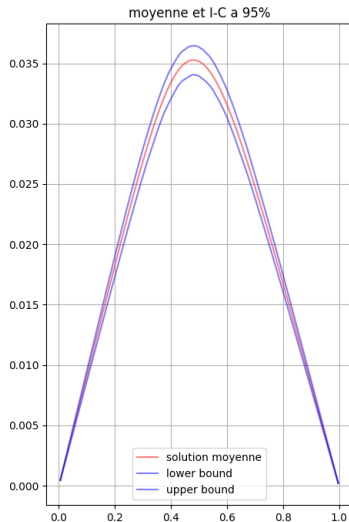
$$\begin{cases} -\frac{\partial^2 u}{\partial x^2} = \theta_1 e^{-\theta_3(x-\theta_2)^2} & \text{pour } x \in (0, 1) \\ u(0) = 0, u(1) = 0 \end{cases}$$

$$\theta_1 \sim \mathcal{N}(1, 0.1), \theta_2 \sim \mathcal{N}(0.5, 0.1), \theta_3 \sim \mathcal{N}(100, 10)$$

On va faire 100 simulation

La courbe rouge représente la solution déterministe, la courbe grise correspond à la solution moyenne, les courbes vertes montrent les percentiles de la distribution des DDLs (ce ne sont donc pas des solutions) et la courbe beige pâle représente la médiane à chaque DDL. Aussi, $COV = 0.02$





On va étudier la déformation d'une poutre d'acier soumise à une force suivant des paramètres aléatoires. On va étudier les statistiques du stress de von Mises (σ_v) moyen. Le stress de von Mises nous apprend que, pour l'acier, lorsque $\sigma_v \geq 0.35$ le matériau est déformé de façon permanente. (110 Simulations)

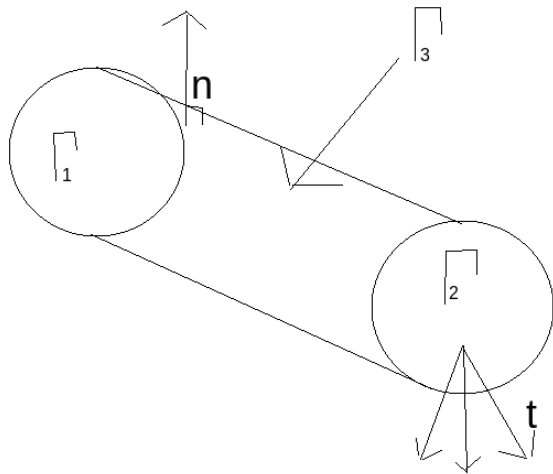
$$\begin{cases} \operatorname{div}(\sigma(u)) = 0 & \text{dans } \Omega \\ u = 0 & \text{sur } \Gamma_1 \\ \sigma(u) \cdot \mathbf{n} = t & \text{sur } \Gamma_2 \\ \sigma(u) \cdot \mathbf{n} & \text{sur } \Gamma_3 \end{cases}$$

$$t = r(\sin(\phi), \sin(\theta), -\cos(\phi) \cos(\theta))$$

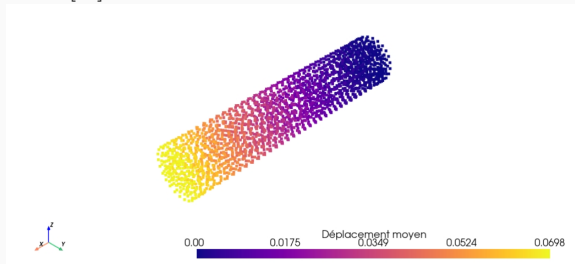
$$r \sim \text{LogNormal}(-5, 1)$$

$$\phi \sim \mathcal{N}(0, \frac{\pi^2}{576})$$

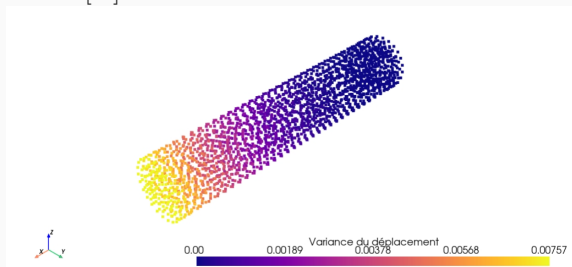
$$\theta \sim \mathcal{N}(0, \frac{\pi^2}{576})$$



$\mathbb{E}^{MC}[U] :$



$\text{Var}^{MC}[U] :$



$$COV = 0.01$$

Formulation : Soit $\theta_{Ddl} = \mathbb{E}[\sigma_v]$ le paramètre d'intérêt, où σ_v est le stress de von Mises à un DDL donné, et $I_\alpha = [T_1, T_2]$ un IC de niveau de confiance $(1 - 2\alpha) \times 100\%$ pour θ_{Ddl} . On a :

- **Si** $0.35 \in I_\alpha$:

Cela indique qu'avec un niveau de confiance de $(1 - 2\alpha) \times 100\%$, la moyenne du stress de von Mises à ce DDL pourrait être supérieur ou inférieur à 0.35. On ne peut donc pas conclure

- **Si** $0.35 \notin I_\alpha$:

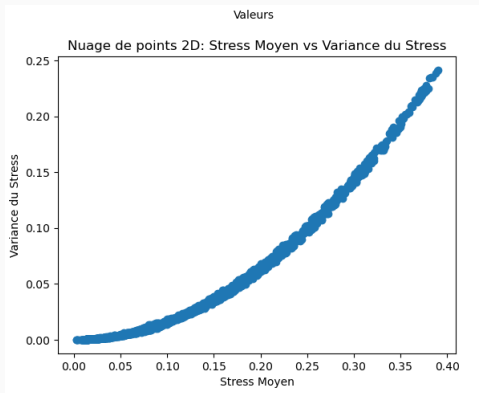
- **Si** $0.35 < T_1$:

Ici, avec un niveau de confiance de $(1 - 2\alpha) \times 100\%$, on conclut que le stress moyen de von Mises à ce DDL est inférieur à 0.35. Par conséquent, la déformation permanente est exclue pour ce DDL.

- **Si** $0.35 > T_2$:

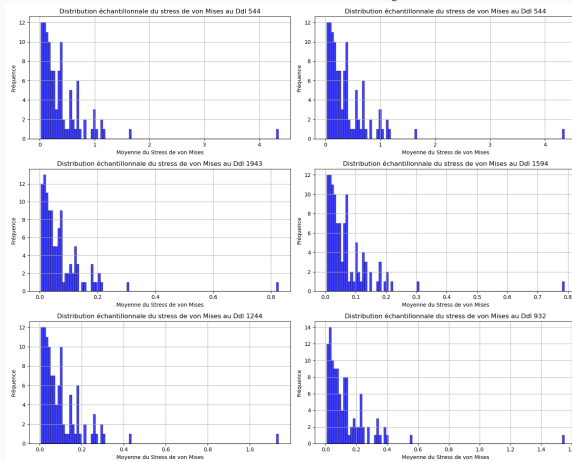
Dans ce cas, avec un niveau de confiance de $(1 - 2\alpha) \times 100\%$, on conclut que le stress moyen de von Mises à ce DDL dépasse 0.35. Ainsi, il y aura déformation permanente en moyenne pour ce DDL.

On veut calculer des I.C sur les données du stress de von Mises a chaque DDL. Pour éviter un nombre élevé de calcul, on prend les DDLs où on a une probabilité significative que 0.35 soit dans l'I.C. Ce graphique liant l'écart-type et la moyenne.



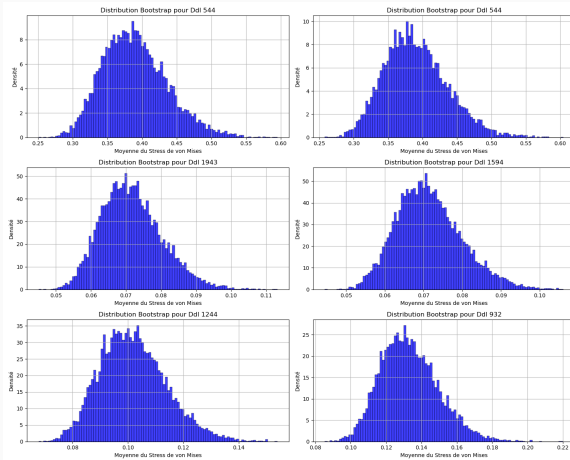
Donc, pour l'expérience, on va seulement étudier les ddl où le stress moyen ≥ 0.2

Distribution échantillonales de σ_v à certain Ddl :



La distribution n'est pas symétrique donc, on peut conclure que la méthode des percentiles ne sera pas efficace.

Distribution Bootstrap approximative pour 10000 échantillons Bootstrap ($B = 10000$) pour les Dddls proche de la médiane , moyenne et valeur maximal pour le stress moyen et la variance du stress

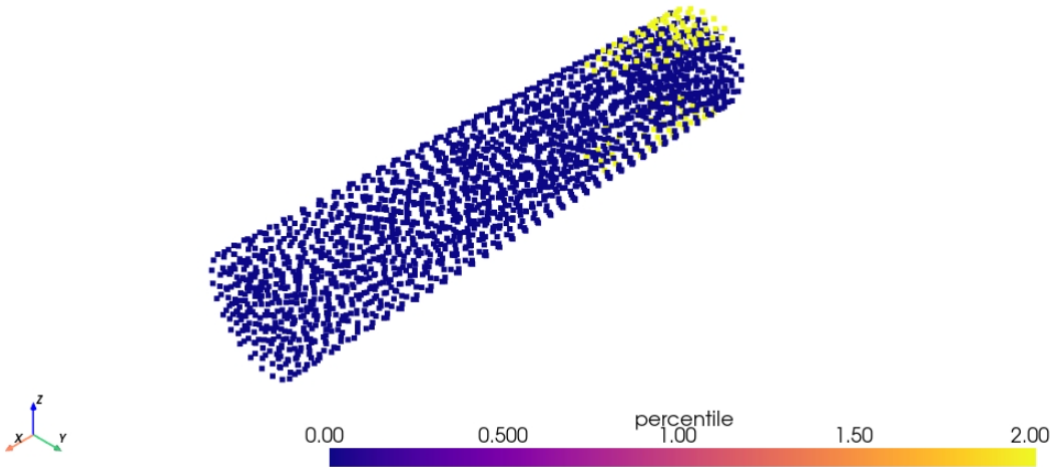


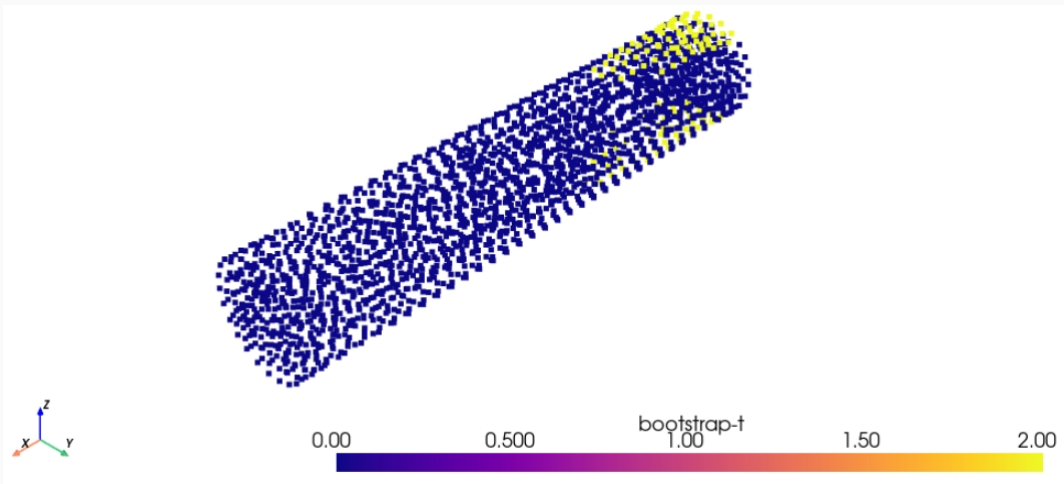
On remarque une légère asymétrie ce qui va impacter l'efficacité du Bootstrap-t

Critères d'interprétation en fonction de la présence de 0.35 dans les intervalles de confiance (I_α) pour le stress moyen de von Mises :

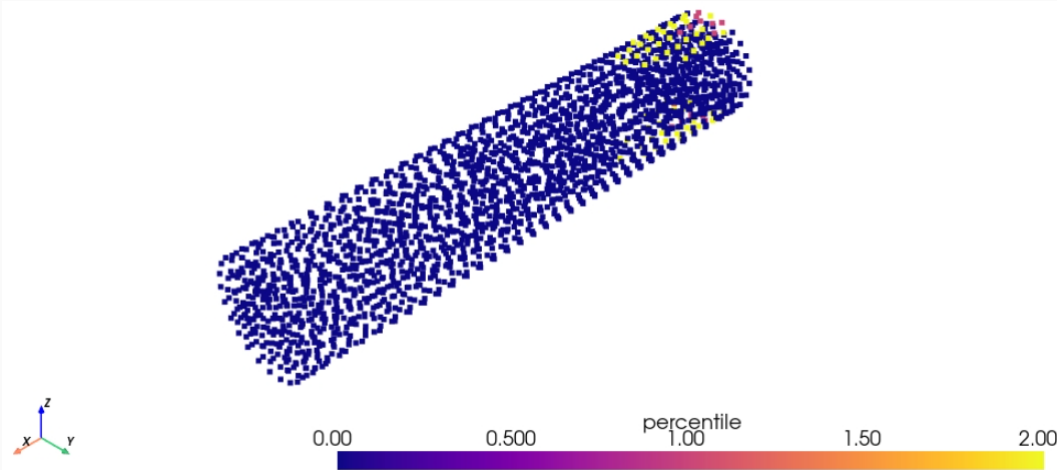
- **Si** $0.35 \in I_\alpha$:
 - Le DDL sera coloré en **jaune**
 - Interprétation : Il existe une incertitude quant à la déformation permanente. Le stress moyen pourrait être supérieur ou inférieur à 0.35.
- **Si** $0.35 < T_1$:
 - Le DDL sera coloré en **bleu**
 - Interprétation : La déformation permanente est exclue, le stress moyen est en dessous de 0.35.
- **Si** $0.35 > T_2$:
 - Le DDL sera coloré en **rose**
 - Interprétation : Le stress moyen est supérieur à 0.35, indiquant une déformation permanente.

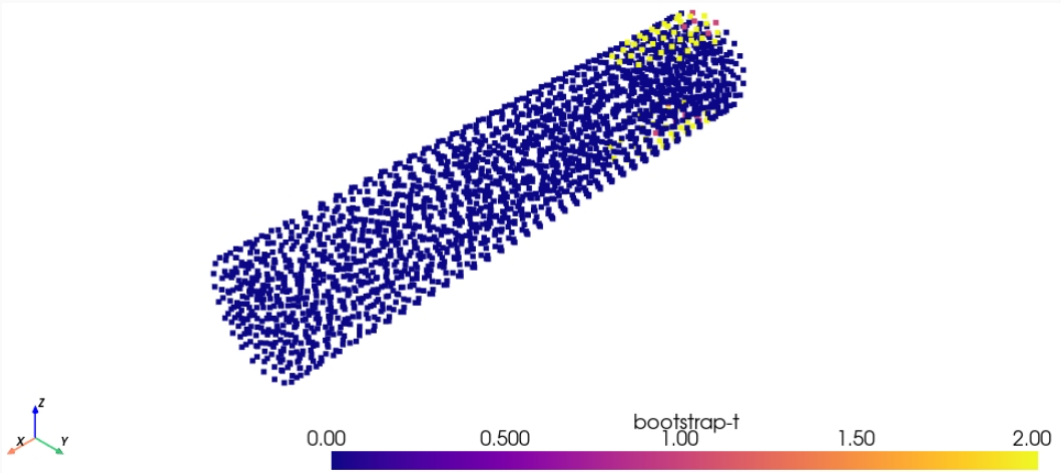
Testé avec un niveau de confiance de 90% :










Testé avec un niveau de confiance de 75% :





-  A. Autor.
Introduction to Giving Presentations.
Klein-Verlag, 1990.
-  J. Shao, D. Tu
The Jackknife and Bootstrap.
Springer, 1995.
-  P. Hall
The Bootstrap and Edgeworth Expansion.
Springer, 1992.
-  B. Surret, A. Der Kiureghian
Stochastic Finite Element Methods and Reliability A State of the Art Report.
Department of civil and environmental engineering university of california , 2000.

 T.J. Sullivan
Introduction to Uncertainty Quantification.
Springer, 2015.

 Jean-Claude Massé
Note de cours stt-66972.
Université Laval, 2009.