

Projet 1 : Simulation et Copules

BARTHE Alexandre - ZULFICAR Eric

15 novembre 2021

On a choisi de travailler sur la fonction :

$$\varphi : (x, y) \mapsto \sqrt{\cosh(x) + \sin^2(y)} \times \log(1 + xy)$$

Voici à quoi ressemble le volume que l'on cherche à estimer sur $[1, 2] \times [1, 3]$.

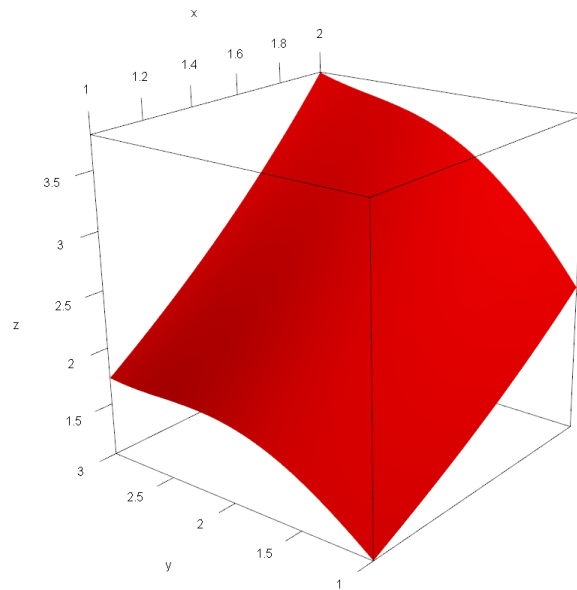


FIGURE 1: Visualisation 3D de notre surface

```
#calcul de l'integrale  
> pracma::quad2d(f = phi, xa = 1,xb = 2, ya = 1, yb = 3)  
[1] 4.7432
```

Après calcul de la double intégrale, on doit donc estimer un volume de 4.7432.

A chaque méthodes, nous avons des échantillons de taille $n = 100000$ et nous effectuons $N = 1000$ simulations de n échantillons.

Méthode 1 : Fréquence empirique

On commence par simuler 3 n-échantillons $X \sim \mathcal{U}_{[1,2]}$, $Y \sim \mathcal{U}_{[1,3]}$ et $Z \sim \mathcal{U}_{[0,\varphi(2,3)]}$ de lois uniformes indépendantes identiquement distribuées qui correspondent aux coordonnées (x, y, z) de nos points tirés aléatoirement dans "le cube" qui englobe notre volume .

Ensuite on calcul la moyenne des points situés à l'intérieur du volume à estimer tout en rejetant ceux qui n'y sont pas en la multipliant par le volume total "du cube". On calcul donc :

$$\underbrace{(3-1)(2-1)K}_{\text{volume du cube}} \times \underbrace{\frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{Z_i \leq \varphi(X_i, Y_i)\}}}_{\text{points dans le volume à estimer}}$$

où $K = \varphi(2, 3)$ est la hauteur maximale du volume à estimer.

Enfin, en appliquant le TCL à notre méthode, on obtient le graphe suivant :

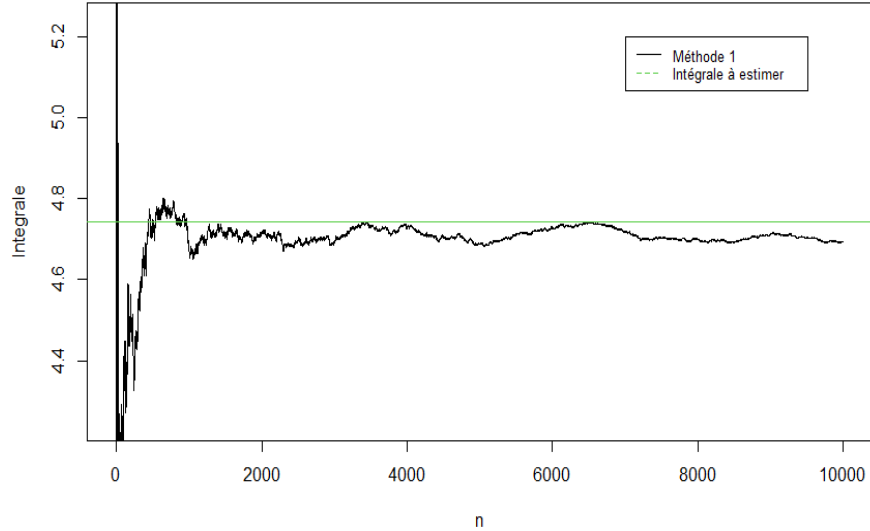


FIGURE 2: Convergence vers le volume de φ sur $[1, 2] \times [1, 3]$

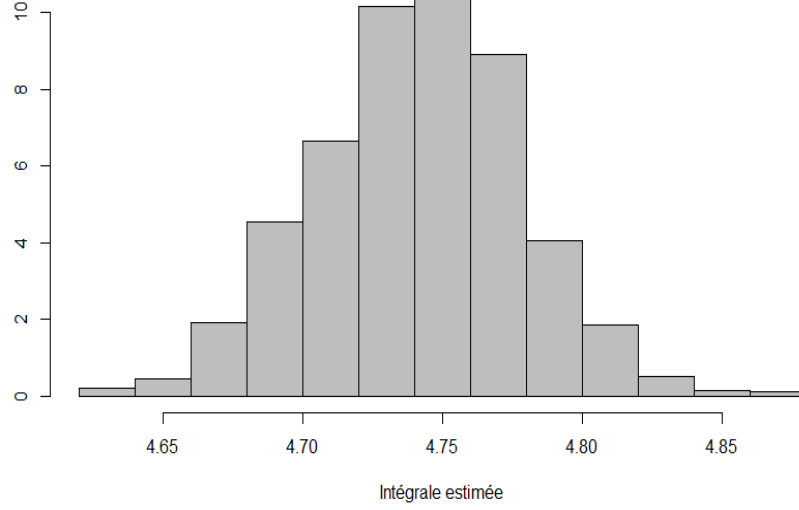


FIGURE 3: Estimation du volume de φ avec 1000 simulations (Méthode 1)

La méthode 1 converge bien vers le volume de φ et sa dispersion n'est pas trop élevée.

Méthode 2 : $\mathcal{U}_{[1,2]} \perp\!\!\!\perp \mathcal{U}_{[1,3]}$

On simule 2 n-échantillons $U = (U_i)_{i=1,\dots,n} \sim \mathcal{U}_{[1,2]}$, $V = (V_i)_{i=1,\dots,n} \sim \mathcal{U}_{[1,3]}$

$$\underbrace{(3-1)(2-1)}_{\text{aire du domaine}} \times \underbrace{\frac{1}{n} \sum_{i=1}^n \varphi(U_i, V_i)}_{\text{valeur moyenne de } \varphi}$$

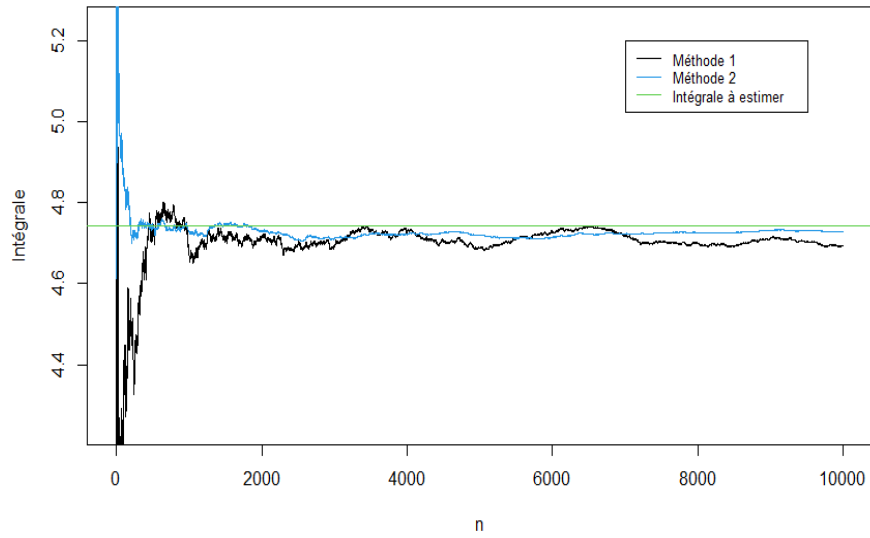


FIGURE 4: Convergence vers le volume de φ sur $[1, 2] \times [1, 3]$

De la même manière, en appliquant le TCL à notre méthode, on obtient le graphe suivant :

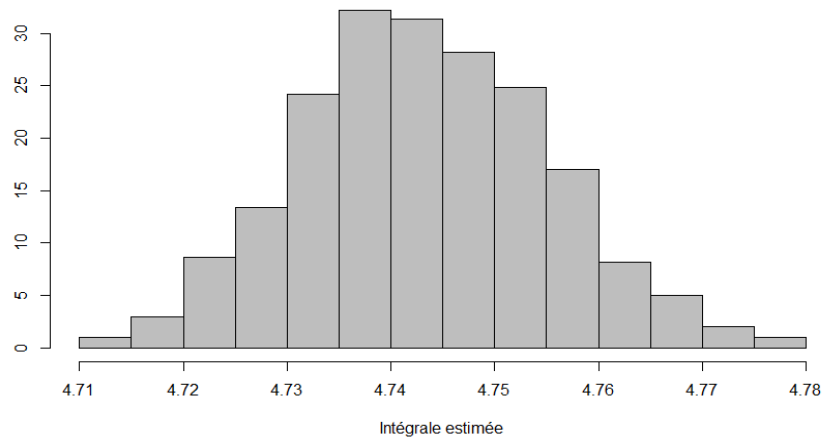


FIGURE 5: Estimation du volume de φ avec 1000 simulations (Méthode 2)

Cette méthode converge légèrement mieux et a une dispersion légèrement moins importante que la première méthode.

Méthode 3 : $p_1(x) \perp\!\!\!\perp p_2(y)$

Pour cette méthode, on va estimer le volume en prenant 2 densités p_1 et p_2 indépendantes telles que, théoriquement, leur produit soit le plus semblable à φ .

On a choisi les densités :

$$p_1(x) = \frac{\cosh(x)}{\sinh(2) - \sinh(1)}, \quad x \in [1, 2]$$

$$p_2(y) = \frac{1}{10}(y + 3), \quad y \in [1, 3]$$

Ce qui donne les fonctions de répartition et leurs inverses suivantes :

$$\left. \begin{array}{l} P_1(x) = \frac{\sinh(x) - \sinh(1)}{c_\alpha} \\ P_1^{-1}(x) = \operatorname{arcsin}(c_\alpha x + \sinh(1)) \\ \text{avec } c_\alpha = \sinh(2) - \sinh(1) \end{array} \right| \begin{array}{l} P_2(x) = \frac{x^2 + 6x - 7}{20} \\ P_2^{-1} = -3 + \sqrt{20(x + \frac{7}{20}) + 9} \end{array}$$

Et l'on va calculer :

$$\mathbb{E}_p \left[\frac{\varphi(X, Y)}{p(X, Y)} \right]$$

où : $p(x, y) = p_1(x)p_2(y)$, $X \sim p_1$ et $Y \sim p_2$

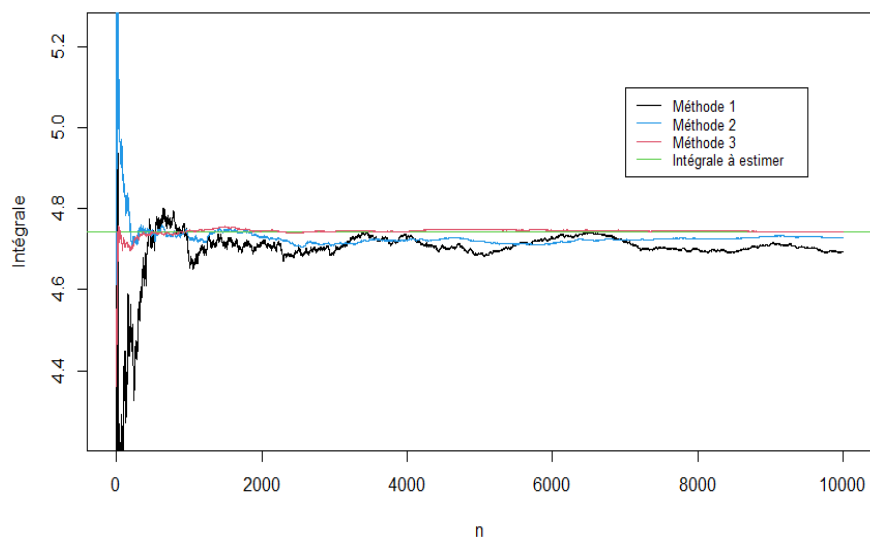


FIGURE 6: Convergence vers le volume de φ sur $[1, 2] \times [1, 3]$

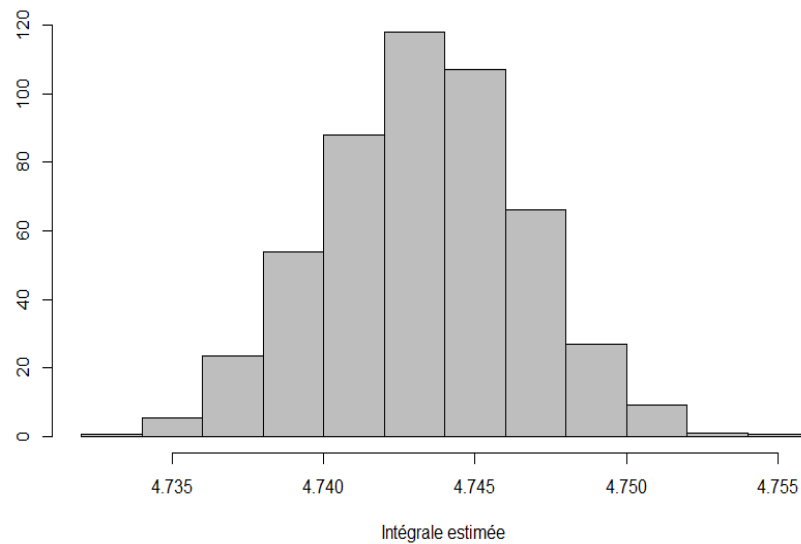


FIGURE 7: Estimation du volume de φ avec 1000 simulations (Méthode 3)

La méthode 3 converge plus rapidement et a une dispersion beaucoup plus faible que les 2 méthodes précédentes.

Méthode 4.1 : Avec la copule de Clayton et des Uniformes sur $[1, 2] \times [1, 3]$

On commence par simuler un couple de variable aléatoire (U_1, V_1) selon la copule de Clayton avec un pic en $(1, 1)$ puis on transforme les marges sur notre domaine.

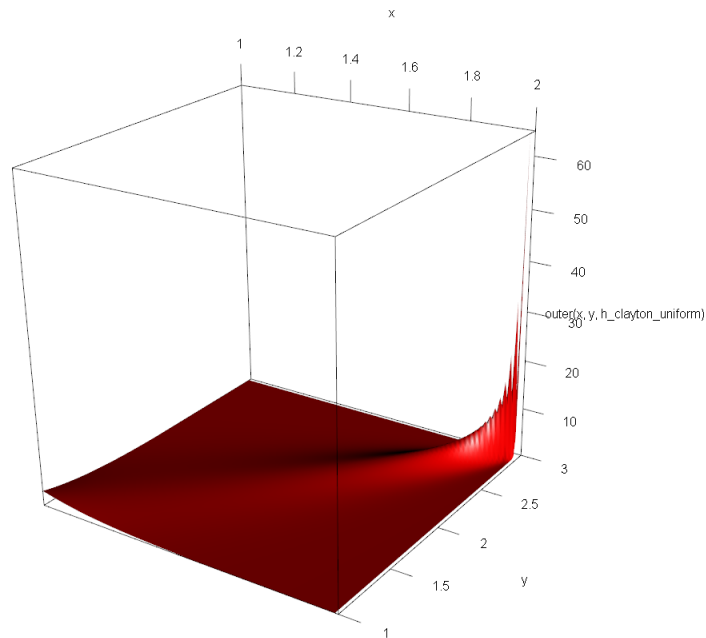


FIGURE 8: Visualisation 3D de la copule de Clayton adaptée à φ

On simule donc le couple : $(F_1^{-1}(U_1), F_2^{-1}(V_1))$ où sont les fonctions de répartitions inverse d'une uniforme $\mathcal{U}[1, 2]$ et $\mathcal{U}[1, 3]$ respectivement.

Ensuite, on va calculer :

$$\frac{1}{n} \sum_{i=1}^n \frac{\varphi(U2, V2)}{p(U2, V2)}$$

où p est obtenu par $p(u, v) = f_1(u)f_2(v) \times c(F_1(u), F_2(v))$ d'après le théorème de Sklar, avec c la densité associé à la copule.

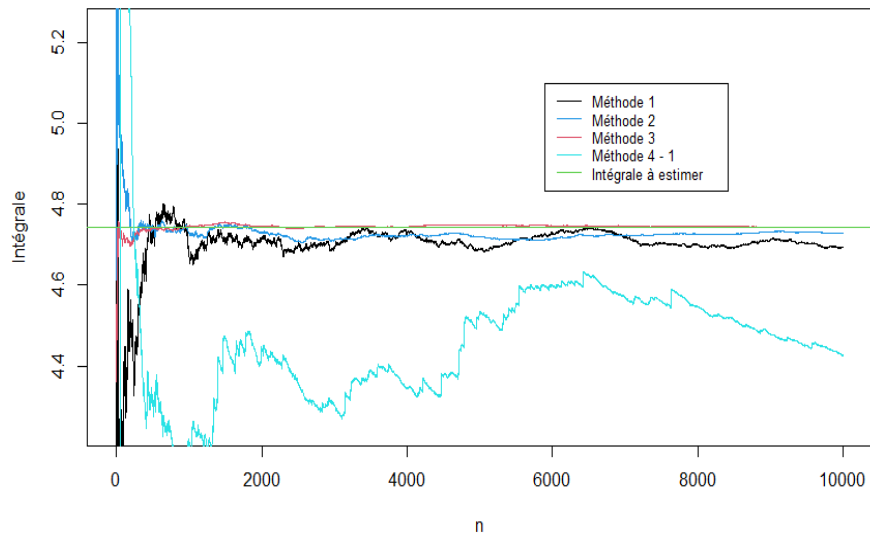


FIGURE 9: Convergence vers le volume de φ sur $[1, 2] \times [1, 3]$

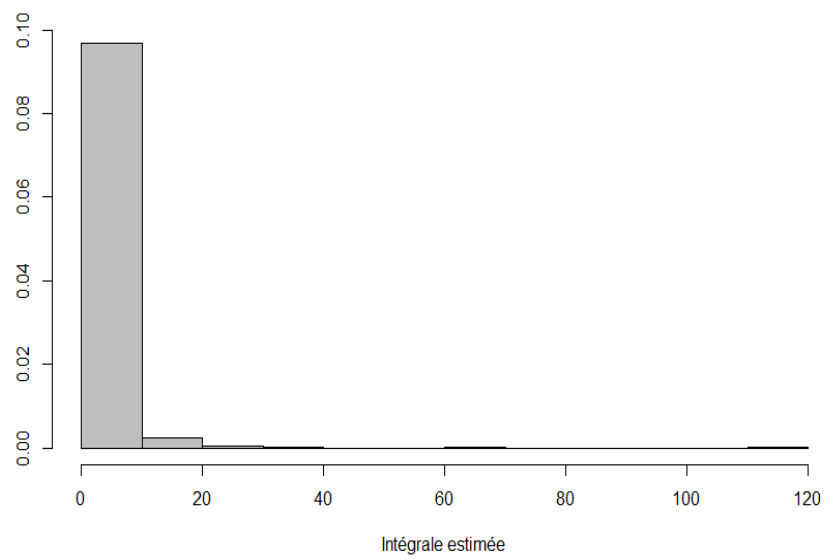


FIGURE 10: Estimation du volume de φ avec 1000 simulations (Méthode 4.1)

La méthode 4.1 (avec $\theta = 2$) converge très lentement vers notre volume à estimer et est très loin de l'efficacité de la méthode 3.

Méthode 4.2 : Avec la copule de Clayton et densités p_1 et p_2

On reprend exactement le même raisonnement que la méthode 4.1 mais au lieu de prendre des densités uniformes et l'inverse de leur fonction de répartition on prend les densités p_1 et p_2 et P_1^{-1} , P_2^{-1} de la méthode 3.

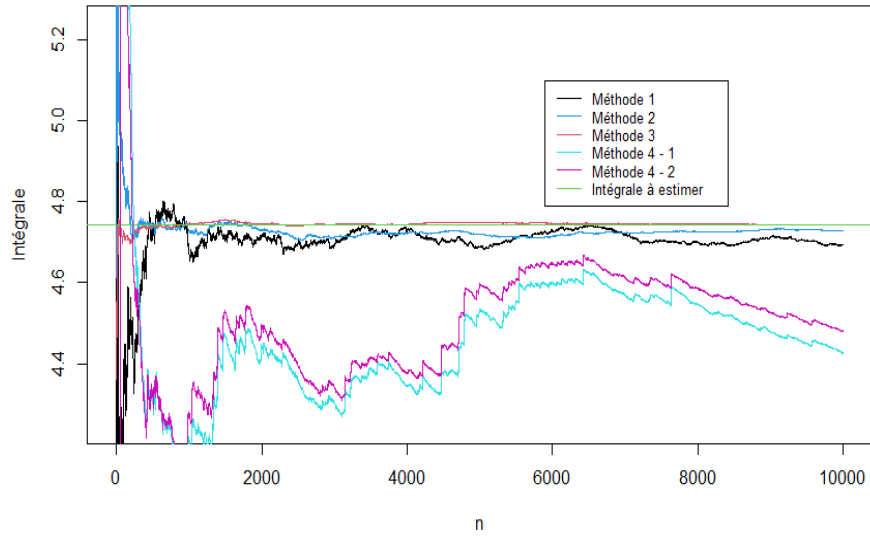


FIGURE 11: Convergence vers le volume de φ sur $[1, 2] \times [1, 3]$

On remarque que la méthode 4.2 (avec $\theta = 2$) converge légèrement mieux que la méthode 4.1 mais estime elle aussi très mal notre volume.

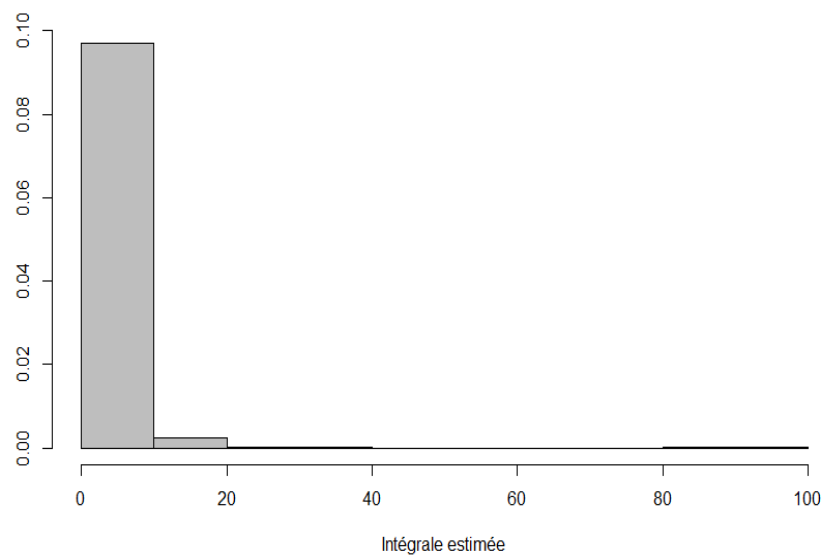


FIGURE 12: Estimation du volume de φ avec 1000 simulations (Méthode 4.2)

Comparaison des méthodes

Pour comparer les méthodes entre elles, on va s'intéresser à leurs écart-type, plus celui-ci est faible, meilleure est la méthode.

```
> sd(TCL_M1) # ecart type methode 1
[1] 0.03659649
> sd(TCL_M2) # ecart type methode 2
[1] 0.01158523
> sd(TCL_M3) # ecart type methode 3
[1] 0.003305548
> sd(TCL_M4_unif) # ecart type methode 4.1
[1] 11.95961
> sd(TCL_M4_m3) # ecart type methode 4.2
[1] 9.739435
```

Ainsi, on obtient que la méthode 3 est, de loin, la plus adaptée pour estimer notre φ . Elle converge très rapidement vers φ et sa dispersion est la plus faible.

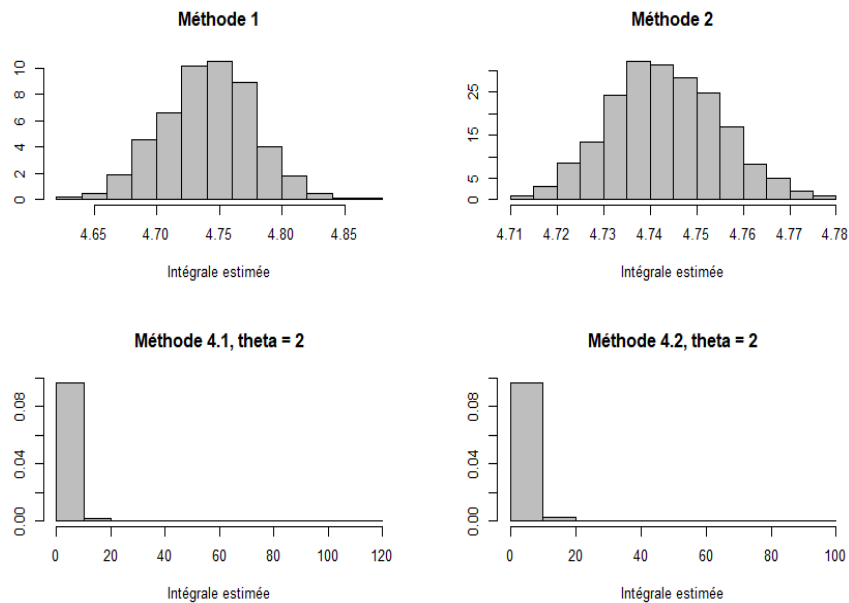


FIGURE 13: Résumé des TCL des méthodes vues précédemment

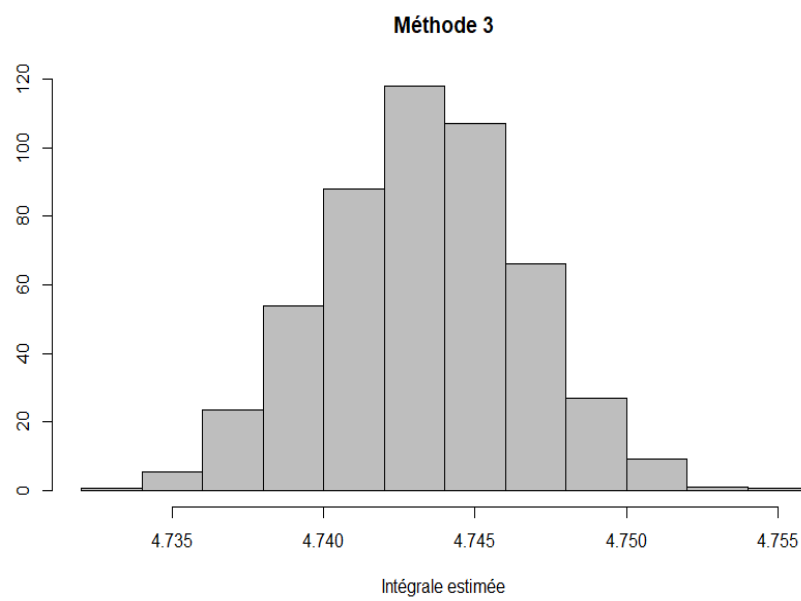


FIGURE 14: Résumé des TCL des méthodes vues précédemment