RÉPUBLIQUE FRANÇAISE
Liberté Égalité Fraternité

Institut Supérieur de l'Aéronautique et de l'Espace SUPAERO

ONERA THE FRENCH AEROSPACE LAB

Control variates for variance reduction in ratio of means estimators

Application to Extreme Value Index estimation

Louison Bocquet-Nouaille[1,2], Jérôme Morio[1], Benjamin Bobbia[2]
[1]ONERA-DTIS, [2]ISAE-DISC

ETICS - 09 octobre 2025

Plan

Goal : Reducing the variance of ratio of means estimators

Goal : estimate a ratio of means

$$R = \frac{\mathbb{E}[A]}{\mathbb{E}[C]}$$

where A and C are arbitrary random variables. Assume $\mathbb{E}[C] \neq 0$.

Goal : Reducing the variance of ratio of means estimators

Goal : estimate a ratio of means

$$R = \frac{\mathbb{E}[A]}{\mathbb{E}[C]}$$

where A and C are arbitrary random variables. Assume $\mathbb{E}[C] \neq 0$.

Classical estimator : ratio of Monte Carlo estimators

$$\widehat{R}_{\frac{MC}{MC}} = \frac{\overline{A_n}}{\overline{C_n}}$$

where $\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i$ the Monte Carlo estimator of $\mathbb{E}[X]$, with $(X_i)_{i=1\dots n}$ i.i.d. samples of a random variable X

Goal : Reducing the variance of ratio of means estimators

Goal : estimate a ratio of means

$$R = \frac{\mathbb{E}[A]}{\mathbb{E}[C]}$$

where A and C are arbitrary random variables. Assume $\mathbb{E}[C] \neq 0$.

Classical estimator : ratio of Monte Carlo estimators

$$\widehat{R}_{\frac{MC}{MC}} = \frac{\overline{A_n}}{\overline{C_n}}$$

where $\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i$ the Monte Carlo estimator of $\mathbb{E}[X]$, with $(X_i)_{i=1\ldots n}$ i.i.d. samples of a random variable X

Problem : high variance of the Monte Carlo estimator

Method : Control variates, applied to a mean estimator

Control variates estimator of $\mathbb{E}[A]$: with control variate B

$$\widehat{A}_{CV} = \overline{A_n} + \alpha_c(\mathbb{E}[B] - \overline{B_n})$$

$(A_i, B_i)_{i=1\dots n}$ i.i.d. samples from the joint distribution of the random variables A and B

Method : Control variates, applied to a mean estimator

Control variates estimator of $\mathbb{E}[A]$ **:** with control variate B

$$\widehat{A}_{CV} = \overline{A_n} + \alpha_c(\mathbb{E}[B] - \overline{B_n})$$

$(A_i, B_i)_{i=1...n}$ i.i.d. samples from the joint distribution of the random variables A and B

$$\alpha_c := \underset{\alpha \in \mathbb{R}}{\text{argmin}} \ \text{Var}\left(\overline{A_n} + \alpha(\mathbb{E}[B] - \overline{B_n})\right) = \frac{\text{Cov}(A, B)}{\text{Var}(B)}$$

$$\text{Var}(\widehat{A}_{CV}) = \left(1 - \text{Corr}(A, B)^2\right) \text{Var}(\overline{A_n})$$

Method : Control variates, applied to a mean estimator

Control variates estimator of $\mathbb{E}[A]$: with control variate B

$$\widehat{A}_{CV} = \overline{A_n} + \alpha_c(\mathbb{E}[B] - \overline{B_n})$$

$(A_i, B_i)_{i=1\dots n}$ i.i.d. samples from the joint distribution of the random variables A and B

$$\alpha_c := \underset{\alpha\in\mathbb{R}}{\operatorname{argmin}} \operatorname{Var}\left(\overline{A_n} + \alpha(\mathbb{E}[B] - \overline{B_n})\right) = \frac{\operatorname{Cov}(A, B)}{\operatorname{Var}(B)}$$

$$\operatorname{Var}(\widehat{A}_{CV}) = \left(1 - \operatorname{Corr}(A, B)^2\right) \operatorname{Var}(\overline{A_n})$$

$\rightarrow$ $|\operatorname{Corr}(A, B)| \nearrow \;\Rightarrow\; \operatorname{Var}(\widehat{A}_{CV}) \searrow$

$\rightarrow$ Worst case scenario $\operatorname{Corr}(A, B) = 0 \Rightarrow \widehat{A}_{CV} = \overline{A_n}$ the Monte-Carlo estimator

Method : Control variates, applied to a ratio of means estimator

Control variates estimator of $R = \frac{\mathbb{E}[A]}{\mathbb{E}[C]}$:

$$\widehat{R}_{\frac{CV}{CV}} = \frac{\widehat{A}_{CV}}{\widehat{C}_{CV}} = \frac{\overline{A_n} + \alpha(\mathbb{E}[B] - \overline{B_n})}{\overline{C_n} + \beta(\mathbb{E}[D] - \overline{D_n})}$$

$(A_i, B_i, C_i, D_i)_{i=1\dots n}$ i.i.d. samples from the joint distribution of the random variables A, B, C, and D

Method : Control variates, applied to a ratio of means estimator

Control variates estimator of $R = \frac{\mathbb{E}[A]}{\mathbb{E}[C]}$:

$$\widehat{R}_{\frac{CV}{CV}} = \frac{\widehat{A}_{CV}}{\widehat{C}_{CV}} = \frac{\overline{A_n} + \alpha(\mathbb{E}[B] - \overline{B_n})}{\overline{C_n} + \beta(\mathbb{E}[D] - \overline{D_n})}$$

- the classical coefficients for mean estimators :

$$\alpha_c := \underset{\alpha \in \mathbb{R}}{\operatorname{argmin}} \operatorname{Var}\left(\overline{A_n} + \alpha(\mathbb{E}[B] - \overline{B_n})\right) \qquad \beta_c := \underset{\beta \in \mathbb{R}}{\operatorname{argmin}} \operatorname{Var}\left(\overline{C_n} + \beta(\mathbb{E}[D] - \overline{D_n})\right)$$

Method : Control variates, applied to a ratio of means estimator

Control variates estimator of $R = \frac{\mathbb{E}[A]}{\mathbb{E}[C]}$:

$$\widehat{R}_{\frac{CV}{CV}} = \frac{\widehat{A}_{CV}}{\widehat{C}_{CV}} = \frac{\overline{A_n} + \alpha(\mathbb{E}[B] - \overline{B_n})}{\overline{C_n} + \beta(\mathbb{E}[D] - \overline{D_n})}$$

- the classical coefficients for mean estimators :

$$\alpha_c := \underset{\alpha\in\mathbb{R}}{\text{argmin}}\ \text{Var}\left(\overline{A_n} + \alpha(\mathbb{E}[B] - \overline{B_n})\right) \qquad \beta_c := \underset{\beta\in\mathbb{R}}{\text{argmin}}\ \text{Var}\left(\overline{C_n} + \beta(\mathbb{E}[D] - \overline{D_n})\right)$$

- the coefficients introduced in [Gordon et al., 1982] :

$$\alpha_g := \underset{\alpha\in\mathbb{R}}{\text{argmin}}\ \text{Var}\left(\overline{A_n} + \alpha(\mathbb{E}[B] - \overline{B_n})\right) \qquad \beta_g := \underset{\beta\in\mathbb{R}}{\text{argmin}}\ \text{Var}\left(\frac{\overline{A_n} + \alpha_g(\mathbb{E}[B] - \overline{B_n})}{\overline{C_n} + \beta(\mathbb{E}[D] - \overline{D_n})}\right)$$

Method : Control variates, applied to a ratio of means estimator

Control variates estimator of $R = \frac{\mathbb{E}[A]}{\mathbb{E}[C]}$ **:**

$$\widehat{R}_{\frac{CV}{CV}} = \frac{\widehat{A}_{CV}}{\widehat{C}_{CV}} = \frac{\overline{A_n} + \alpha(\mathbb{E}[B] - \overline{B_n})}{\overline{C_n} + \beta(\mathbb{E}[D] - \overline{D_n})}$$

- the classical coefficients for mean estimators :

$$\alpha_c := \underset{\alpha \in \mathbb{R}}{\text{argmin}} \text{ Var}\left(\overline{A_n} + \alpha(\mathbb{E}[B] - \overline{B_n})\right) \qquad \beta_c := \underset{\beta \in \mathbb{R}}{\text{argmin}} \text{ Var}\left(\overline{C_n} + \beta(\mathbb{E}[D] - \overline{D_n})\right)$$

- the coefficients introduced in [Gordon et al., 1982] :

$$\alpha_g := \underset{\alpha \in \mathbb{R}}{\text{argmin}} \text{ Var}\left(\overline{A_n} + \alpha(\mathbb{E}[B] - \overline{B_n})\right) \qquad \beta_g := \underset{\beta \in \mathbb{R}}{\text{argmin}} \text{ Var}\left(\frac{\overline{A_n} + \alpha_g(\mathbb{E}[B] - \overline{B_n})}{\overline{C_n} + \beta(\mathbb{E}[D] - \overline{D_n})}\right)$$

$\rightarrow$ with these coefficients, the variance can be increased with certain covariance structures

Academic use case : Illustration of the variance increase

Ratio to estimate $R = \frac{\mathbb{E}[A]}{\mathbb{E}[C]}$
with control variates B and D

$n = 100$ samples
$(A, B, C, D)^\top \sim \mathcal{N}(\mu, \Sigma)$

$$\mu = \begin{pmatrix} 50 \\ 20 \\ 10 \\ 100 \end{pmatrix} \quad \Sigma = \begin{pmatrix} 1 & & & \\ -0.99 & 1 & & \\ 0.99 & -0.99 & 1 & \\ -0.02 & 0.02 & -0.01 & 1 \end{pmatrix}$$

Reported simulation values averaged over
10,000 independent repetitions.

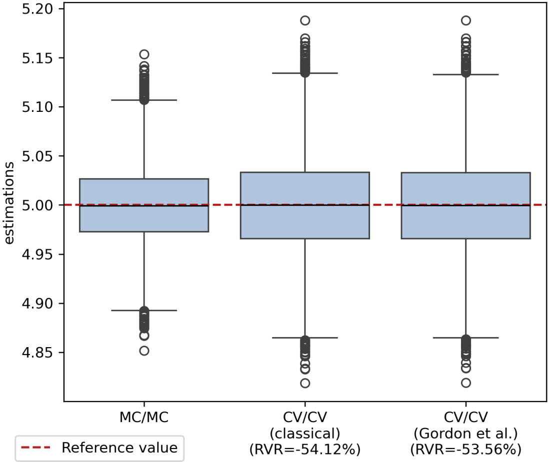
Academic use case : Illustration of the variance increase

Ratio to estimate $R = \frac{\mathbb{E}[A]}{\mathbb{E}[C]}$
with control variates B and D

$n = 100$ samples
$(A, B, C, D)^\top \sim \mathcal{N}(\mu, \Sigma)$

$$\mu = \begin{pmatrix} 50 \\ 20 \\ 10 \\ 100 \end{pmatrix} \quad \Sigma = \begin{pmatrix} 1 & & & \\ -0.99 & 1 & & \\ 0.99 & -0.99 & 1 & \\ -0.02 & 0.02 & -0.01 & 1 \end{pmatrix}$$

Reported simulation values averaged over 10,000 independent repetitions.



Proposed method : Control variates with jointly optimized coefficients

Control variates estimator of $R = \frac{\mathbb{E}[A]}{\mathbb{E}[C]}$ **:** $\widehat{R}_{\frac{CV}{CV}} = \frac{\widehat{A}_{CV}}{\widehat{C}_{CV}} = \frac{\overline{A_n}+\alpha(\mathbb{E}[B]-\overline{B_n})}{\overline{C_n}+\beta(\mathbb{E}[D]-\overline{D_n})}$

- proposed optimal coefficients :

$$(\alpha_o, \beta_o) := \underset{(\alpha,\beta)\in\mathbb{R}^2}{\text{argmin}} \, \text{Var}\left(\frac{\overline{A_n} + \alpha(\mathbb{E}[B] - \overline{B_n})}{\overline{C_n} + \beta(\mathbb{E}[D] - \overline{D_n})}\right)$$

Proposed method : Control variates with jointly optimized coefficients

Control variates estimator of $R = \frac{\mathbb{E}[A]}{\mathbb{E}[C]}$ **:** $\widehat{R}_{\frac{CV}{CV}} = \frac{\widehat{A}_{CV}}{\widehat{C}_{CV}} = \frac{\overline{A_n}+\alpha(\mathbb{E}[B]-\overline{B_n})}{\overline{C_n}+\beta(\mathbb{E}[D]-\overline{D_n})}$

- proposed optimal coefficients :

$$(\alpha_o, \beta_o) := \underset{(\alpha,\beta)\in\mathbb{R}^2}{\text{argmin}} \text{ Var}\left(\frac{\overline{A_n} + \alpha(\mathbb{E}[B] - \overline{B_n})}{\overline{C_n} + \beta(\mathbb{E}[D] - \overline{D_n})}\right)$$

$$\alpha_o = \frac{\text{Var}(D)\text{Cov}(A, B) - R\text{Var}(D)\text{Cov}(B, C) + R\text{Cov}(B, D)\text{Cov}(C, D) - \text{Cov}(B, D)\text{Cov}(A, D)}{\text{Var}(B)\text{Var}(D) - \text{Cov}(B, D)^2}$$

$$\beta_o = \frac{\text{Cov}(B, D)\text{Cov}(A, B) - R\text{Cov}(B, D)\text{Cov}(B, C) + R\text{Var}(B)\text{Cov}(C, D) - \text{Var}(B)\text{Cov}(A, D)}{R\left(\text{Var}(B)\text{Var}(D) - \text{Cov}(B, D)^2\right)}$$

Proposed method : Control variates with jointly optimized coefficients

Control variates estimator of $R = \frac{\mathbb{E}[A]}{\mathbb{E}[C]}$ **:** $\widehat{R}_{CV} = \frac{\widehat{A}_{CV}}{\widehat{C}_{CV}} = \frac{\overline{A_n}+\alpha(\mathbb{E}[B]-\overline{B_n})}{\overline{C_n}+\beta(\mathbb{E}[D]-\overline{D_n})}$

- proposed optimal coefficients :

$$(\alpha_o, \beta_o) := \underset{(\alpha,\beta)\in\mathbb{R}^2}{\text{argmin}} \ \text{Var}\left(\frac{\overline{A_n} + \alpha(\mathbb{E}[B] - \overline{B_n})}{\overline{C_n} + \beta(\mathbb{E}[D] - \overline{D_n})}\right)$$

$$\alpha_o = \frac{\text{Var}(D)\text{Cov}(A,B) - R\text{Var}(D)\text{Cov}(B,C) + R\text{Cov}(B,D)\text{Cov}(C,D) - \text{Cov}(B,D)\text{Cov}(A,D)}{\text{Var}(B)\text{Var}(D) - \text{Cov}(B,D)^2}$$

$$\beta_o = \frac{\text{Cov}(B,D)\text{Cov}(A,B) - R\text{Cov}(B,D)\text{Cov}(B,C) + R\text{Var}(B)\text{Cov}(C,D) - \text{Var}(B)\text{Cov}(A,D)}{R\left(\text{Var}(B)\text{Var}(D) - \text{Cov}(B,D)^2\right)}$$

$\rightarrow$ with these coefficients, the variance reduction is **theoretically guaranteed**

if $|\text{Corr}(B,D)| < 1$

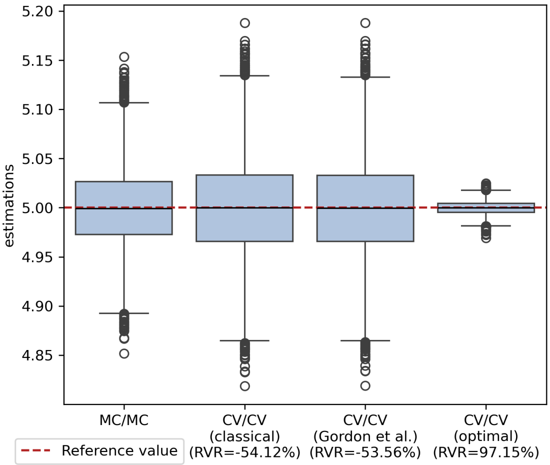
Academic use case : Illustration of the variance reduction guarantee

Ratio to estimate $R = \frac{\mathbb{E}[A]}{\mathbb{E}[C]}$
with control variates B and D

$n = 100$ samples
$(A, B, C, D)^\top \sim \mathcal{N}(\mu, \Sigma)$

$$\Sigma = \begin{pmatrix} 1 & & & \\ -0.99 & 1 & & \\ 0.99 & -0.99 & 1 & \\ -0.02 & 0.02 & -0.01 & 1 \end{pmatrix}$$

Reported simulation values averaged over 10,000 independent repetitions.



Aircraft design application : Illustration of the variance reduction guarantee

Ratio to estimate : strut mass / total mass

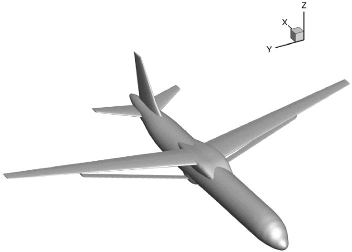


Figure – ONERA ALBATROS configuration with strut-braced wings (Figure 11 from [Carrier et al., 2022])

$$\text{Correlation matrix} = \begin{pmatrix} 1 & & & \\ 0.51 & 1 & & \\ 0.77 & 0.4 & 1 & \\ 0.83 & 0.78 & 0.74 & 1 \end{pmatrix}$$

Aircraft design application : Illustration of the variance reduction guarantee

Ratio to estimate : strut mass / total mass

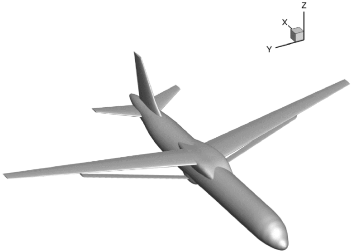


Figure – ONERA ALBATROS configuration with strut-braced wings (Figure 11 from [Carrier et al., 2022])

$$\text{Correlation matrix} = \begin{pmatrix} 1 & & & \\ 0.51 & 1 & & \\ 0.77 & 0.4 & 1 & \\ 0.83 & 0.78 & 0.74 & 1 \end{pmatrix}$$

Multi-fidelity dataset of 1252 samples ($n = 200$)

Aircraft design application : Illustration of the variance reduction guarantee

Ratio to estimate : strut mass / total mass

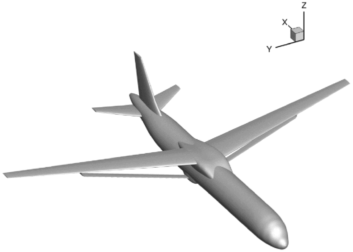
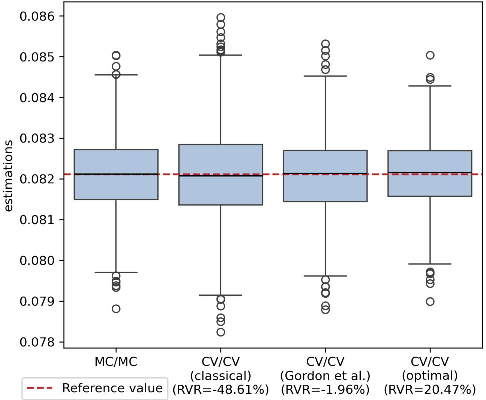


Figure – ONERA ALBATROS configuration with strut-braced wings (Figure 11 from [Carrier et al., 2022])

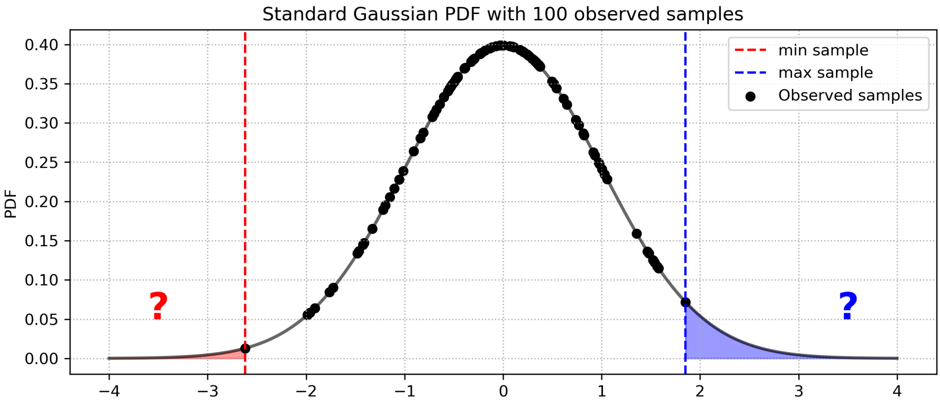
$$\text{Correlation matrix} = \begin{pmatrix} 1 & & & \\ 0.51 & 1 & & \\ 0.77 & 0.4 & 1 & \\ 0.83 & 0.78 & 0.74 & 1 \end{pmatrix}$$

Multi-fidelity dataset of 1252 samples ($n = 200$)



A (very) brief introduction to Extreme Value Theory

[Bousquet and Bernardara, 2021][de Haan and Ferreira, 2006]



γ **the Extreme Value Index**, key parameter characterizing tails of probability distributions

A variance-reduced Hill estimator of the Extreme Value Index

Hill estimator (Monte-Carlo ratio estimator) $\widehat{\gamma}_H^T = \frac{\overline{A_n}}{\overline{C_n}}$

$$A = (\ln(Y^T) - \ln(Y^T_{n-k:n}))\mathbb{1}_{\{Y^T > Y^T_{n-k:n}\}} \qquad C = \mathbb{1}_{\{Y^T > Y^T_{n-k:n}\}}$$

(target)

A variance-reduced Hill estimator of the Extreme Value Index

Hill estimator (Monte-Carlo ratio estimator) $\widehat{\gamma}_H^T = \frac{\overline{A_n}}{\overline{C_n}}$

$$A = (\ln(Y^T) - \ln(Y^T_{n-k:n}))\mathbb{1}_{\{Y^T > Y^T_{n-k:n}\}} \qquad C = \mathbb{1}_{\{Y^T > Y^T_{n-k:n}\}} \qquad \text{(target)}$$

$$B = (\ln(Y^S) - \ln(Y^S_{n-k:n}))\mathbb{1}_{\{Y^S > Y^S_{n-k:n}\}} \qquad D = \mathbb{1}_{\{Y^S > Y^S_{n-k:n}\}} \qquad \text{(source)}$$

Variance-reduced Hill estimator with **approximate** control variates

$$\widehat{\gamma}_{TH}^T = \frac{\overline{A_n} + \alpha(\overline{B_{n+m}} - \overline{B_n})}{\overline{C_n} + \beta(\overline{D_{n+m}} - \overline{D_n})}$$

with $k \in \{1, \ldots, n-1\}$, $Y^T_{n-k:n} > 0$, and $|\mathrm{Corr}(B, D)| < 1$

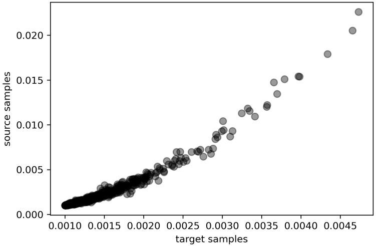
Results with the variance-reduced Hill estimator

Pareto marginals

$\gamma^T = 0.25, \gamma^S = 0.5$

$n = 1,000, \; k = 100, \; m = 5,000$

Gumbel copula with parameter $\theta = 10$



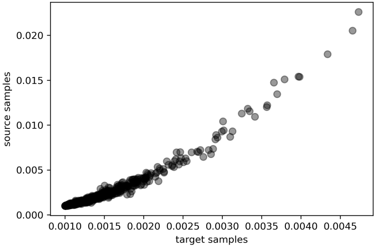
Results with the variance-reduced Hill estimator

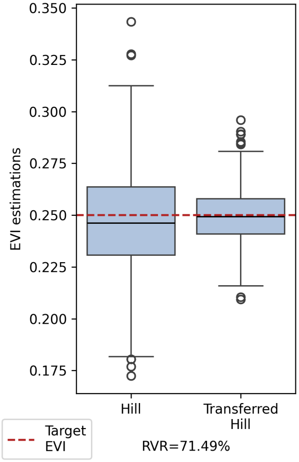
Pareto marginals

$\gamma^T = 0.25, \gamma^S = 0.5$

$n = 1,000,\ k = 100,\ m = 5,000$

Gumbel copula with parameter $\theta = 10$





Conclusion

Take-home message

Control variates coefficients need to be optimized for the whole estimator
to keep the variance reduction guarantee

Conclusion

Take-home message

Control variates coefficients need to be optimized for the whole estimator to keep the variance reduction guarantee

Code available at `https://github.com/LouB-N`

Conclusion

Take-home message

Control variates coefficients need to be optimized for the whole estimator
to keep the variance reduction guarantee

Code available at `https://github.com/LouB-N`

Thank you for your attention.
Questions ?

Bibliographie

[Bousquet and Bernardara, 2021] Bousquet, N. and Bernardara, P. (2021).
Extreme Value Theory with Applications to Natural Hazards From Statistical Theory to Industrial Practice.

[Carrier et al., 2022] Carrier, G. G., Arnoult, G., Fabbiane, N., Schotte, J.-S., David, C., Defoort, S., Benard, E., and Delavenne, M. (2022).
Multidisciplinary analysis and design of strut-braced wing concept for medium range aircraft.
In *AIAA SCITECH 2022 Forum*, pages –.

[de Haan and Ferreira, 2006] de Haan, L. and Ferreira, A. (2006).
Extreme Value Theory : An Introduction.

[Gordon et al., 1982] Gordon, H. L., Rothstein, S. M., and Proctor, T. R. (1982).
Efficient variance-reduction transformations for the simulation of a ratio of two means : application to quantum Monte Carlo simulations.
Journal of Computational Physics, 47(3) :375–386.

Derivating the optimal coefficients

$$R = \frac{\mathbb{E}[A]}{\mathbb{E}[C]} \qquad \widehat{R}_{\frac{MC}{MC}} = \frac{\overline{A_n}}{\overline{C_n}} \qquad \widehat{R}_{\frac{CV}{CV}} = \frac{\widehat{A}_{CV}}{\widehat{C}_{CV}} = \frac{\overline{A_n} + \alpha(\mathbb{E}[B] - \overline{B_n})}{\overline{C_n} + \beta(\mathbb{E}[D] - \overline{D_n})}$$

$$Var\left(\frac{\hat{A}_{CV,n}}{\hat{C}_{CV,n}}\right) = \frac{1}{\mathbb{E}[\hat{C}_{CV,n}]^2}\mathrm{Var}(\hat{A}_{CV}) + \frac{\mathbb{E}[\hat{A}_{CV}]^2}{\mathbb{E}[\hat{C}_{CV,n}]^4}\mathrm{Var}(\hat{C}_{CV,n}) - 2\frac{\mathbb{E}[\hat{A}_{CV}]}{\mathbb{E}[\hat{C}_{CV,n}]^3}\mathrm{Cov}\left(\hat{A}_{CV}, \hat{C}_{CV,n}\right)$$

$$\nabla Var\left(\frac{\hat{A}_{CV,n}}{\hat{C}_{CV,n}}\right) = \begin{bmatrix}0\\0\end{bmatrix} \iff M \cdot \begin{bmatrix}\alpha\\ \beta\end{bmatrix} = \begin{bmatrix}c_1\\c_2\end{bmatrix}$$

$$\Leftrightarrow \alpha = \frac{\mathrm{Var}(D)\mathrm{Cov}(A,B) - R\mathrm{Var}(D)\mathrm{Cov}(B,C) + R\mathrm{Cov}(B,D)\mathrm{Cov}(C,D) - \mathrm{Cov}(B,D)\mathrm{Cov}(A,D)}{\mathrm{Var}(B)\mathrm{Var}(D) - \mathrm{Cov}(B,D)^2}$$

$$\beta = \frac{\mathrm{Cov}(B,D)\mathrm{Cov}(A,B) - R\mathrm{Cov}(B,D)\mathrm{Cov}(B,C) + R\mathrm{Var}(B)\mathrm{Cov}(C,D) - \mathrm{Var}(B)\mathrm{Cov}(A,D)}{R\left(\mathrm{Var}(B)\mathrm{Var}(D) - \mathrm{Cov}(B,D)^2\right)}$$

M inversible if $|\mathrm{Corr}(B, D)| < 1$

We show that $\mathrm{Var}\left(\widehat{R}_{\frac{CV}{CV}}\right) \leq \mathrm{Var}\left(\widehat{R}_{\frac{MC}{MC}}\right)$.

Optimal coefficients for the CV/CV estimator if $|\text{Corr}(B, D)| = 1$

Assume $\forall (a, b) \in \mathbb{R}^* \times \mathbb{R}, B = aD + b$

$$\begin{cases} \alpha_o = \frac{1}{a}\left(\frac{\text{Cov}(A,D) - R\text{Cov}(C,D)}{\text{Var}(D)} + \beta_o R\right) \\ \beta_o \in \mathbb{R} \end{cases}$$

Variance reduction if

$$\text{Cov}(A - RC, D) \in \left(-\sqrt{\frac{\text{Var}(D)}{2}}, 0\right) \cup \left(\sqrt{\frac{\text{Var}(D)}{2}}, +\infty\right)$$

L'indice des valeurs extrêmes (EVI) γ

$\rightarrow$ **Approche BM "block maxima"** [de Haan and Ferreira, 2006] [Bousquet and Bernardara, 2021]

Théorème (Loi limite du maximum) : Soit F la fonction de répartition de la v.a.r. Y.
$\exists(a_n, b_n)$ deux suites telles que $a_n > 0$ et

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{max(Y_1, \ldots, Y_n) - b_n}{a_n} \leq y\right) = G_\gamma(y)$$

où G_γ non dégénérée* $\Rightarrow F \in DA(G_\gamma)$ (F dans le domaine d'attraction de G_γ)

$$G_\gamma(y) = \begin{cases} exp\left(-(1+\gamma y)^{-\frac{1}{\gamma}}\right) & \text{si } \gamma \neq 0, \quad 1+\gamma y > 0 \\ exp\left(-e^{-y}\right) & \text{si } \gamma = 0 \end{cases}$$

(Loi d'Extremum Généralisée)

γ **indice des valeurs extrêmes (EVI)** $\rightarrow$ caractérise la queue de distribution

* non dégénérée = pas presque sûrement constante = de variance > 0

L'indice des valeurs extrêmes (EVI) γ

$\rightarrow$ **Approche POT "peaks over threshold"** [de Haan and Ferreira, 2006] [Bousquet and Bernardara, 2021]

Théorème (Loi limite des excès de seuil) :

- $F_u(y) = \mathbb{P}(Y - u > y | Y > u)$ la fonction de répartition des excès de seuil,
- y^* le point terminal défini par $y^* = sup\{y : F(y) < 1\}$,
- $\sigma(u)$ une fonction de normalisation positive.

$$F \in DA(G_\gamma) \Leftrightarrow \lim_{u \to y^*} F_u(\sigma(u)y) = GP_\gamma(y)$$

$$GP_\gamma(y) = \begin{cases} 1 - (1 + \gamma y)^{-\frac{1}{\gamma}} & \text{si } \gamma \neq 0, \quad 1 + \gamma y > 0 \\ 1 - exp(-y) & \text{si } \gamma = 0. \end{cases}$$

(Loi de Pareto Généralisée)

γ **indice des valeurs extrêmes (EVI)** $\rightarrow$ caractérise la queue de distribution

L'indice des valeurs extrêmes (EVI)

$\gamma > 0$: queue lourde, point terminal y^* infini $\rightarrow$ **extrêmes relativement fréquents**
Applications : catastrophes naturelles, assurance, finance, cyber-attaques
Lois usuelles : Pareto, Student

$\gamma < 0$: queue courte, point terminal y^* fini
Applications : quantités bornées comme les durées de vie ou des quantités physiques
Lois usuelles : Uniforme, Beta

$\gamma = 0$: queue légère, point terminal y^* fini ou infini $\rightarrow$ **extrêmes relativement rares**
Applications : temps d'attentes
Lois usuelles : Normale, Exponentielle, Gamma, Weibull

y^* le point terminal défini par $y^* = sup\{y : F(y) < 1\}$. [de Haan and Ferreira, 2006] [Bousquet and Bernardara, 2021]