Monte Carlo Method

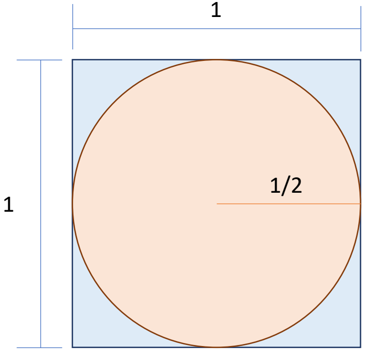
Prof. Americo Cunha Jr

americocunha.org

@AmericoCunhaJr

@AmericoCunhaJr @AmericoCunhaJr

How to estimate the area of a ◯ inscribed in an unit □ ?



$A_{\square} = 1$

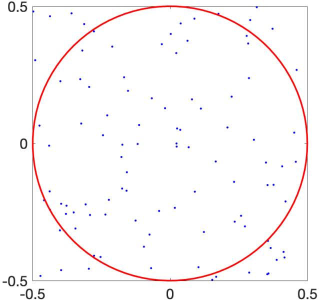
$A_{\bigcirc} = \pi/4$

Recipe to estimate $A_{\bigcirc}$:

1. Draw random pairs: (X, Y) (defined on the unit □)
2. Count how many pairs (X, Y) are inside the ◯
3. An estimate for $A_{\bigcirc}$ is given by the ratio between the pairs within ◯ and the total of pairs

How to estimate the area of a $\bigcirc$ inscribed in an unit $\square$?

$N_s = 100$



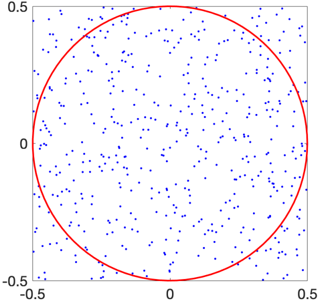
$$\pi/4 = 0.7854\cdots$$

$$\widehat{A}_{\bigcirc} = \frac{75}{100} = 0.7500$$

$$\widehat{\pi} = 3.0000$$

How to estimate the area of a $\bigcirc$ inscribed in an unit $\square$?

$N_s = 500$



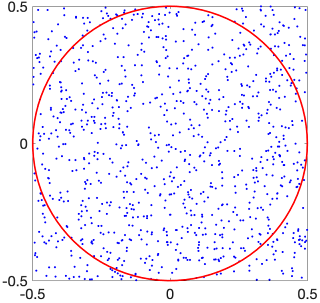
$$\pi/4 = 0.7854\cdots$$

$$\widehat{A}_{\bigcirc} = \frac{395}{500} = 0.7900$$

$$\widehat{\pi} = 3.1600$$

How to estimate the area of a ◯ inscribed in an unit □ ?

$N_s = 1000$



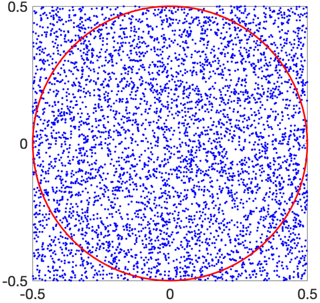
$$\pi/4 = 0.7854\cdots$$

$$\widehat{A}_{\bigcirc} = \frac{794}{1000} = 0.7940$$

$$\widehat{\pi} = 3.1760$$

How to estimate the area of a $\bigcirc$ inscribed in an unit $\square$?

$N_s = 5000$



$$\pi/4 = 0.7854\cdots$$

$$\widehat{A}_{\bigcirc} = \frac{3960}{5000} = 0.7920$$

$$\widehat{\pi} = 3.1680$$

Overview of Monte Carlo (MC) Method

Monte Carlo Method

pre-processing	**processing**	**post-processing**
generation of scenarios	solution of model equations	computation of statistics
$\mathbf{X}_1$		$Y_1 = \mathcal{M}(\mathbf{X}_1)$
$\vdots$	**computational model**	$\vdots$
$\mathbf{X}_n$		$Y_n = \mathcal{M}(\mathbf{X}_n)$
known $F_{\mathbf{X}}$		estimated F_Y
generator of random vector $\mathbf{X}$	deterministic solver of $y = \mathcal{M}(\mathbf{x})$	statistical inference to estimate convergence and distribution of Y

Monte Carlo as an estimation technique

Generaly, in Monte Carlo method, one wants to compute

$$\mu = \mathbb{E}\left\{Y\right\} = \mathbb{E}\left\{\mathcal{M}(\mathbf{X})\right\} = \int_{\mathbb{R}} \mathcal{M}(\mathbf{x})\, p_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x}$$

An estimator for such integral is given by

$$\widehat{\mu} = \frac{1}{n} \sum_{j=1}^{n} \mathcal{M}(\mathbf{X}_j),$$

where $\mathbf{X}_1, \cdots, \mathbf{X}_n$ are independent observations of $\mathbf{X}$.

C. Soize, *Uncertainty Quantification: An Accelerated Course with Advanced Applications in Computational Engineering*, Springer, 2017.

Properties of the estimator

- unbiased

$$\mathbb{E}\left\{\widehat{\mu}\right\} = \mu$$

- mean-square error (variance) proportional to $1/n$

$$\texttt{MSE}\left(\widehat{\mu}\right) = \sigma^2_{\widehat{\mu}} = \frac{1}{n}\,\sigma^2_Y$$

- convergence guaranteed when Y is second-ordered

$$\texttt{MSE}\left(\widehat{\mu}\right) \to 0 \;\; \text{when} \;\; n \to +\infty \quad \Longleftrightarrow \quad \mathbb{E}\left\{Y^2\right\} < +\infty$$

C. Soize, *Uncertainty Quantification: An Accelerated Course with Advanced Applications in Computational Engineering*, Springer, 2017.

Monte Carlo features

Good features:

- ☺ easy implementation
- ☺ non-intrusive solver
- ☺ embarrassingly parallel algorithm
- ☺ convergence guaranteed
  (law of large numbers)
- ☺ convergence controled by number of realizations
  (central limit theorem)
- ☺ convergence independent of dimension
  (no curse of dimensionality)

Bad features:

- ☹ slow convergence $\sim 1/\sqrt{n}$
- ☹ high computational cost

MC method may be used for:

- Calculation of hyper-volumes
- Numerical integration
- Statistical estimation
  - mean, variance, skewness, kurtosis
  - other statistical moments
  - confidence intervals
  - probabilities of interest
  - etc
- Propagation of uncertainties
- many other (probabilistic or deterministic) tasks

MC for numerical integration

Consider the definite integral

$$\alpha = \int_{\mathbb{R}} f(x)\, dx.$$

Note that

$$\int_{\mathbb{R}} f(x)\, dx = \int_{\mathbb{R}} \frac{f(y)}{p(y)}\, p(y)\, dy = \int_{\mathbb{R}} h(y)\, p(y)\, dy = \mathbb{E}\left\{h(Y)\right\},$$

where Y is a random variable with density $p(y)$ and $h(y) = f(y)/p(y)$. Thus, it is possible to estimate α by

$$\widehat{\alpha} = \frac{1}{n} \sum_{j=1}^{n} h(Y_j),$$

where $Y_1, \cdots, Y_n$ are independent observations of Y.

MC for numerical integration

Consider the integral

$$\begin{aligned}\alpha &= \int_0^{2\pi} \sin^2(x)\, dx \\ &= \int_0^{2\pi} 2\pi \sin^2(x)\, \frac{1}{2\pi}\, dx \\ &= \int_{\mathbb{R}} h(y)\, p(y) dy\end{aligned}$$

where

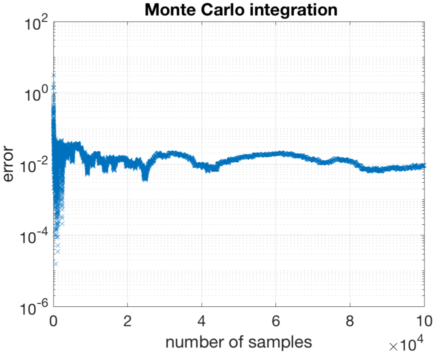
- $h(y) = 2\pi \sin^2(y)$
- $p(y) = \frac{1}{2\pi} \mathbb{1}_{(0,2\pi)}(y)$ i.e., $Y \sim \mathcal{U}(0, 2\pi)$.

MC for numerical integration

Tabela: Evaluation of the integral via Monte Carlo Method.

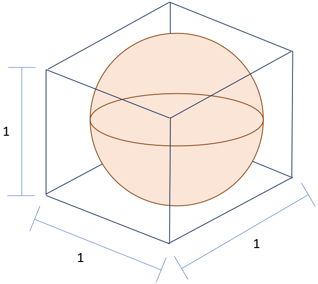
samples	$\widehat{\alpha}$	$\lvert\alpha - \widehat{\alpha}\rvert$ **
10	3.7543	3.1392
100	3.2420	0.1004
1000	3.1349	0.0066
10000	3.1256	0.0159
100000	3.1324	0.0091

**Analytical solution $\alpha = \pi$



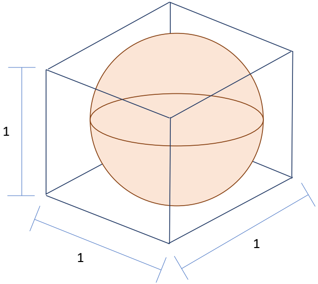
MC for estimation of statistics

What is the probability that you randomly draw a point inside a unit cube and it is contained in the inscribed sphere?



MC for estimation of statistics

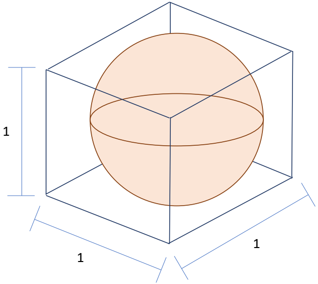
What is the probability that you randomly draw a point inside a unit cube and it is contained in the inscribed sphere?



$$\mathcal{P}\left\{(X, Y, Z) \in x^2 + y^2 + z^3 \leq 1/4\right\} = \frac{1}{6}\,\pi$$

MC for estimation of statistics

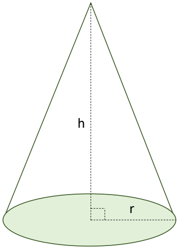
What is the probability that you randomly draw a point inside a unit cube and it is contained in the inscribed sphere?



samples	$\widehat{\mathcal{P}}$	relative error
100	0.5000	4.5%
1000	0.5180	1.1%
10000	0.5319	1.6%
100000	0.5224	0.2%

$$\mathcal{P}\left\{(X, Y, Z) \in x^2 + y^2 + z^3 \leq 1/4\right\} = \frac{1}{6}\,\pi$$

MC for uncertainty propagation



- Parameters support:

$$(r, h) \in [0, 1] \times (0, \infty)$$

- Volume:

$$v(r, h) = \frac{1}{3} \pi \, r^2 \, h$$

- Input joint-distribution:

$R \sim Beta(a, b)$ and $H \sim Exp(1/\mu_H)$ are independent RVs

$$p_{RH}\,(r, h) = \mathbb{1}_{[0,1]}(r) \, \frac{\Gamma\,(a+b)}{\Gamma\,(a)\,\Gamma\,(b)} r^{a-1}\,(1-r)^{b-1} \times \mathbb{1}_{(0,\infty)}(h) \frac{\exp\left(-h/\mu_H\right)}{\mu_H}$$

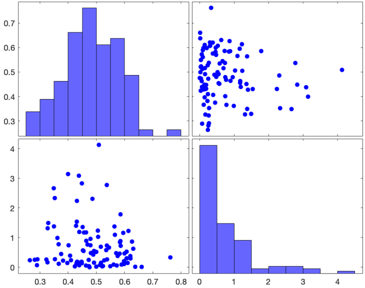
- Output distribution:

$$p_V\,(v) = \,???$$

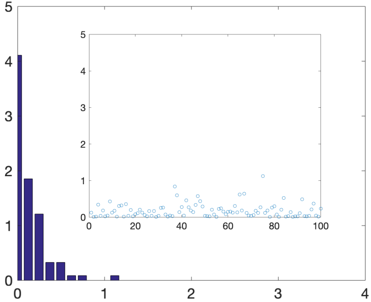
MC for uncertainty propagation

$$N_s = 100$$

Input: $p_{RH}(r, h)$



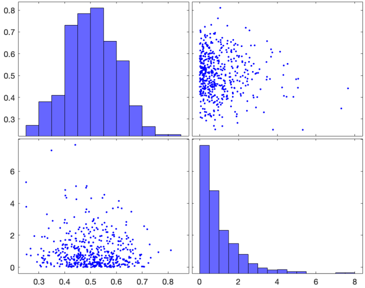
Output: $p_V(v)$



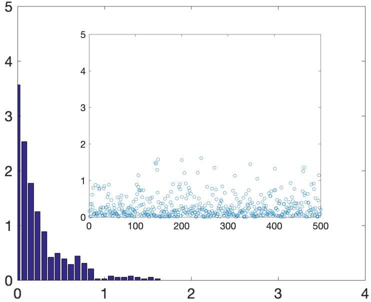
MC for uncertainty propagation

$$N_s = 500$$

Input: $p_{RH}(r, h)$



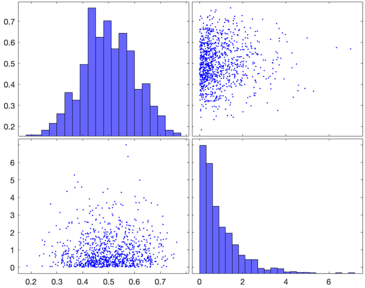
Output: $p_V(v)$



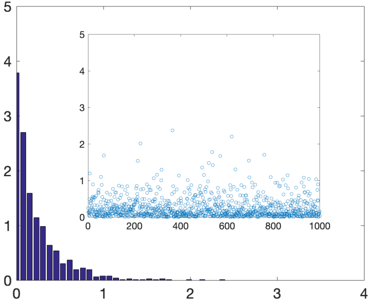
MC for uncertainty propagation

$$N_s = 1000$$

Input: $p_{RH}(r, h)$



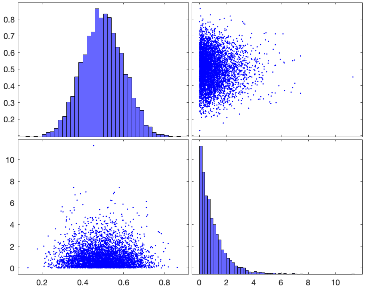
Output: $p_V(v)$



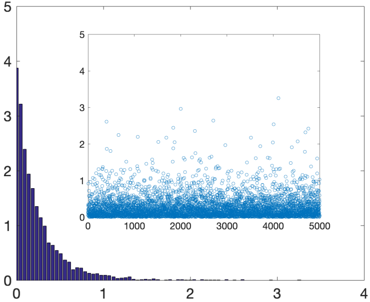
MC for uncertainty propagation

$$N_s = 5000$$

Input: $p_{RH}(r, h)$



Output: $p_V(v)$



An experiment with two different random variables

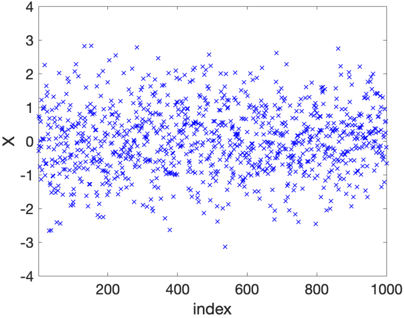
Let's simulate the following random variables:

$$Y = \frac{1}{1 + X^2} \text{ where } X \sim \mathcal{N}(0, 1)$$

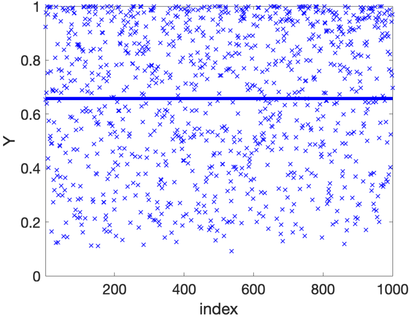
$$Z = \frac{1}{1 + X} \quad \text{where } X \sim \mathcal{N}(0, 1)$$

1000 samples for Y

$$X \sim \mathcal{N}(0,1)$$



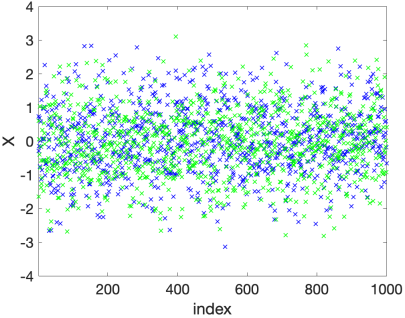
$$Y = \frac{1}{1+X^2}$$



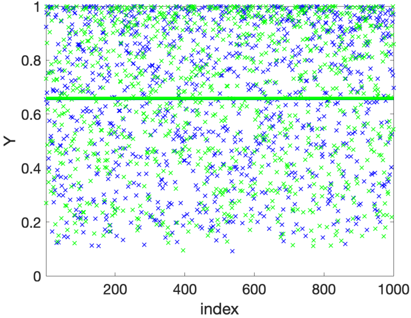
$$\widehat{\mu_Y} = 0.6577$$

1000 samples for Y

$$X \sim \mathcal{N}(0,1)$$



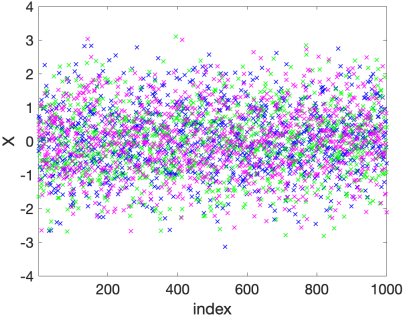
$$Y = \frac{1}{1+X^2}$$



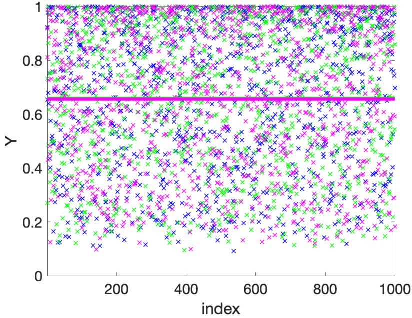
$$\widehat{\mu_Y} = 0.6585$$

1000 samples for Y

$$X \sim \mathcal{N}(0,1)$$



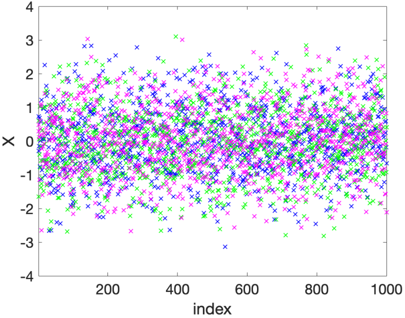
$$Y = \frac{1}{1+X^2}$$



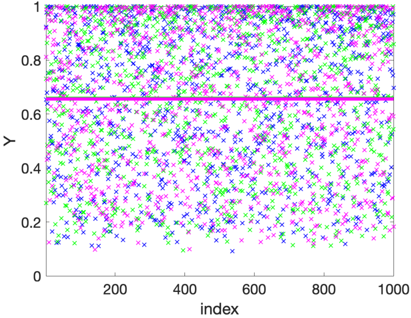
$$\widehat{\mu_Y} = 0.6560$$

1000 samples for Y

$$X \sim \mathcal{N}(0,1)$$



$$Y = \frac{1}{1+X^2}$$

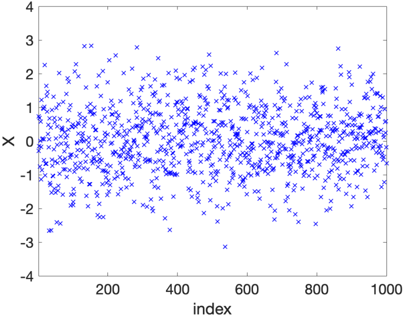


$$\widehat{\mu_Y} = 0.6560$$

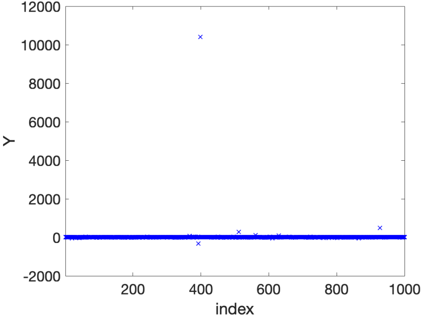
The mean estimation is robust! ☺

1000 samples for Z

$$X \sim \mathcal{N}(0,1)$$



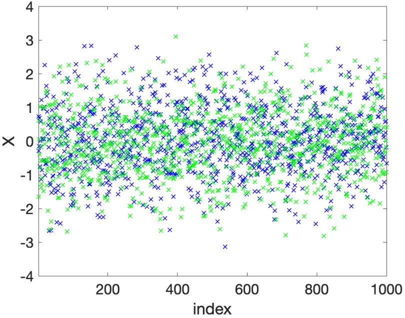
$$Z = \frac{1}{1+X}$$



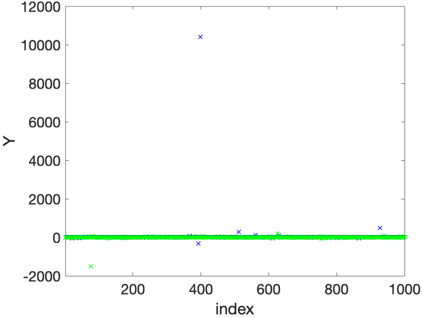
$$\widehat{\mu_Z} = 12.0274$$

1000 samples for Z

$$X \sim \mathcal{N}(0,1)$$



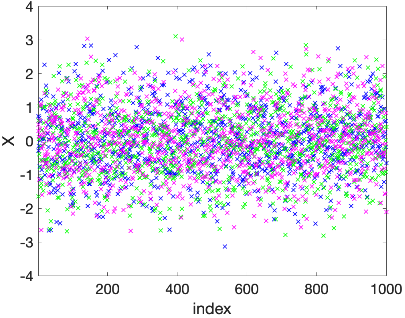
$$Z = \frac{1}{1+X}$$



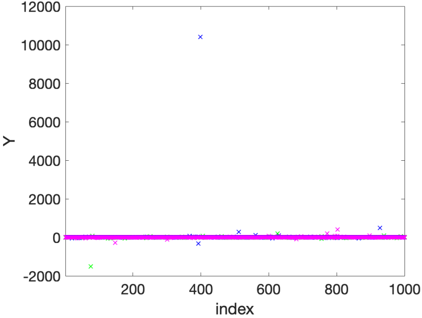
$$\widehat{\mu_Z} = -0.3833$$

1000 samples for Z

$$X \sim \mathcal{N}(0,1)$$



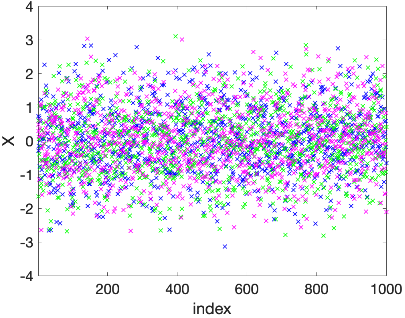
$$Z = \frac{1}{1+X}$$



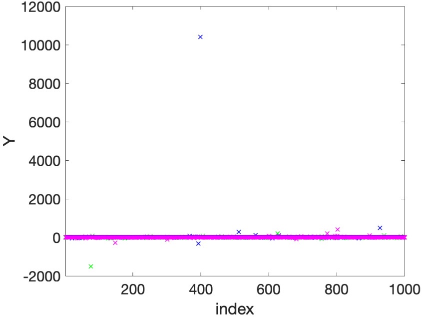
$$\widehat{\mu_Z} = 1.1738$$

1000 samples for Z

$$X \sim \mathcal{N}(0, 1)$$



$$Z = \frac{1}{1 + X}$$



$$\widehat{\mu_Z} = 1.1738$$

The mean estimation is not robust! ☹

Why is this behavior observed?

- $Y = \dfrac{1}{1 + X^2}$ with $X \sim \mathcal{N}(0,1)$

- $Z = \dfrac{1}{1 + X}$ with $X \sim \mathcal{N}(0,1)$

Why is this behavior observed?

- $Y = \dfrac{1}{1+X^2}$ with $X \sim \mathcal{N}(0,1)$

$$\mathbb{E}\left\{Y\right\} = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \left(\frac{1}{1+x^2}\right) \exp\left(-\frac{x^2}{2}\right)\, dx = 0.65568\cdots$$

Y is a finite mean random variable

- $Z = \dfrac{1}{1+X}$ with $X \sim \mathcal{N}(0,1)$

Why is this behavior observed?

- $Y = \dfrac{1}{1+X^2}$ with $X \sim \mathcal{N}(0,1)$

$$\mathbb{E}\left\{Y\right\} = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \left(\frac{1}{1+x^2}\right) \exp\left(-\frac{x^2}{2}\right)\, dx = 0.65568\cdots$$

Y is a finite mean random variable

- $Z = \dfrac{1}{1+X}$ with $X \sim \mathcal{N}(0,1)$

$$\mathbb{E}\left\{Z\right\} = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \left(\frac{1}{1+x}\right) \exp\left(-\frac{x^2}{2}\right)\, dx = \infty$$

Z is an infinite mean random variable
(Monte Carlo simulation does not converge in this case)

Analysis of Monte Carlo convergence

Law of large numbers ensures the convergence of the MC simulation of the random variable Y whenever

$$\mathbb{E}\left\{ Y^2 \right\} < \infty \, ,$$

and the last condition also ensures convergence in distribution:

$$Y_n \xrightarrow{m.s.} Y \implies Y_n \xrightarrow{d} Y \, .$$

In practical terms, an estimation of the second-order moment of Y, computed with aid of the estimator

$$\widehat{\alpha}_2 = \frac{1}{n} \sum_{j=1}^{n} Y_j^2 \, ,$$

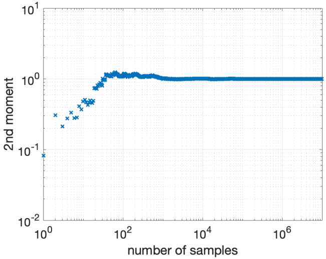
is used to study the MC simulation convergence.

Analysis of Monte Carlo convergence

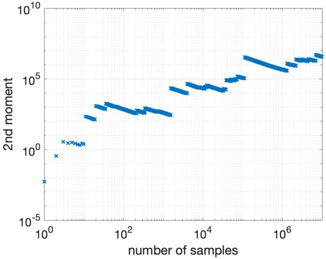
$$\widehat{\alpha}_2 = \frac{1}{n} \sum_{j=1}^{n} Y_j^2$$

(2nd moment estimator)

$Y \sim \mathcal{N}(0,1)$



$Y \sim Cauchy(0,1)$



Analysis of Monte Carlo convergence

Why is this behavior observed?

- $Y \sim \mathcal{N}(0, 1)$

- $Y \sim Cauchy(0, 1)$

Analysis of Monte Carlo convergence

Why is this behavior observed?

- $Y \sim \mathcal{N}(0,1)$

$$\mathbb{E}\left\{Y^2\right\} = \int_{-\infty}^{+\infty} \frac{y^2}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) dy = 1 < \infty$$

- $Y \sim Cauchy(0,1)$

Analysis of Monte Carlo convergence

Why is this behavior observed?

- $Y \sim \mathcal{N}(0,1)$

$$\mathbb{E}\left\{Y^2\right\} = \int_{-\infty}^{+\infty} \frac{y^2}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) dy = 1 < \infty$$

- $Y \sim Cauchy(0,1)$

$$\mathbb{E}\left\{Y^2\right\} \propto \int_{-\infty}^{+\infty} \frac{y^2}{1+y^2} dx = \int_{-\infty}^{+\infty} dy - \pi = \infty$$

References

S. B. Volchan, *What Is a Random Sequence?* **The American Mathematical Monthly**, 109: 46–63, 2002.

T. Homem-de-Mello and G. Bayraksan, *Monte Carlo sampling-based methods for stochastic optimization*, **Surveys in Operations Research and Management Science**, 19: 56–85, 2014.

A. Cunha Jr, R. Sampaio, R. Nasser, H. Lopes, and K. Breitman, *Uncertainty quantification through Monte Carlo method in a cloud computing setting*, **Computer Physics Communications**, 185: 1355-1363, 2014.

R. W. Shonkwiler, and F. Mendivil, **Explorations in Monte Carlo Methods**, Springer, 2009.

C. Robert, and G. Casella, **Monte Carlo Statistical Methods**, Springer, 2nd edition, 2005.

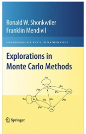
S. Asmussen, and P. W. Glynn, **Stochastic Simulation: Algorithms and Analysis**, Springer, 2007.

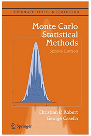
J. E. Gentle, **Random Number Generation and Monte Carlo Methods**, Springer, 2nd Edition, 2004.

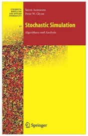
J. S. Liu, **Monte Carlo Strategies in Scientific Computing**, Springer, 2009.

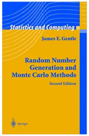
R. Y. Rubinstein, and Dirk P. Kroese, **Simulation and the Monte Carlo Method**, Wiley, 3rd Edition, 2016.

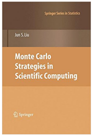
D. P. Kroese, T. Taimre, and Z. I. Botev, **Handbook of Monte Carlo Methods**, Wiley, 2011.

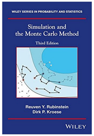


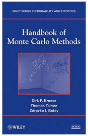












How to cite this material?

A. Cunha Jr, *Monte Carlo Method*, 2021.

@AmericoCunhaJr

@AmericoCunhaJr @AmericoCunhaJr