MATH-414 – Stochastic simulation

Lecture 7: Variance Reduction Techniques III

Prof. Fabio Nobile

Outline

Control variates

Goal: compute $\mu = \mathbb{E}[Z]$ with Z output of a stochastic model

Idea: find an auxiliary random variable Y s.t.

- $\mathbb{E}[Y]$ known
- Y is informative on Z (highly correlated with Z)

Then, for $\alpha \in \mathbb{R}$ set

$$Z_\alpha = Z - \alpha(Y - \mathbb{E}[Y])$$

and apply Monte Carlo on Z_α instead of Z

$$\hat{\mu}_{CV} = \frac{1}{N} \sum_{i=1}^{N} \left[Z^{(i)} - \alpha(Y^{(i)} - \mathbb{E}[Y]) \right], \qquad (Z^{(i)}, Y^{(i)}) \overset{iid}{\sim} (Z, Y)$$

Clearly $\mathbb{E}[Z_\alpha] = \mathbb{E}[Z] = \mu$ so $\hat{\mu}_N^{CV}$ is unbiased.

Do we get varianve reduction?

Control variates - variance minimization

$\mathbb{V}\mathrm{ar}\left[\hat{\mu}_{CV}\right] = \frac{\mathbb{V}\mathrm{ar}[Z_\alpha]}{N}$ and

$$\mathbb{V}\mathrm{ar}\left[Z_\alpha\right] = \mathbb{V}\mathrm{ar}\left[Z\right] + \alpha^2 \mathbb{V}\mathrm{ar}\left[Y\right] - 2\alpha\, \mathrm{Cov}(Z, Y).$$

We can then look for the best α_{opt} that minimizes the variance

$$\frac{d\mathbb{V}\mathrm{ar}\left[Z_\alpha\right]}{d\alpha} = 2\alpha \mathbb{V}\mathrm{ar}\left[Y\right] - 2\,\mathrm{Cov}(Z, Y) = 0 \qquad \Longrightarrow \qquad \alpha_{\mathrm{opt}} = \frac{\mathrm{Cov}(Z, Y)}{\mathbb{V}\mathrm{ar}\left[Y\right]}$$

Minimal variance achievable

$$\mathbb{V}\mathrm{ar}\left[Z_{\alpha_{\mathrm{opt}}}\right] = \mathbb{V}\mathrm{ar}\left[Z\right] - \frac{\mathrm{Cov}(Z, Y)^2}{\mathbb{V}\mathrm{ar}\left[Y\right]} = \mathbb{V}\mathrm{ar}\left[Z\right](1 - \rho_{ZY}^2)$$

with $\rho_{ZY}^2 = \frac{\mathrm{Cov}(Z,Y)^2}{\mathbb{V}\mathrm{ar}[Z]\mathbb{V}\mathrm{ar}[Y]}$ correlation between Z and Y.

$\Longrightarrow$ $\mathbb{V}\mathrm{ar}\left[Z_{\alpha_{\mathrm{opt}}}\right] \leq \mathbb{V}\mathrm{ar}\left[Z\right]$ and variance reduction increases as $\rho_{ZY} \to \{-1, 1\}$. Ideal case: $Y = \pm\gamma Z$ for which $\rho_{ZY} = \pm 1$ and $\mathbb{V}\mathrm{ar}\left[Z_{\alpha_{opt}}\right] = 0$!

Useless since $\mathbb{E}\left[Y\right] = \gamma \mathbb{E}\left[Z\right]$ unknown. But Y should resemble Z.

Control variates – algorithm with pilot run

In practice, α_{opt} non known but can be estimated from a pilot run

---

Algorithm: Control variate with pilot run.

---

1 Generate $\bar{N}$ iid replicas $(Z^{(i)}, Y^{(i)})$, $i = 1, \dots, \bar{N}$ of (Z, Y)

2 Estimate $\hat{\alpha}_{\text{opt}} = \frac{\hat{\sigma}^2_{ZY}}{\sigma^2_Y}$ if σ^2_Y known, or $\hat{\alpha}_{\text{opt}} = \frac{\hat{\sigma}^2_{ZY}}{\hat{\sigma}^2_Y}$ otherwise, with

$$\hat{\sigma}^2_{ZY} = \frac{1}{\bar{N}-1}\sum_{i=1}^{\bar{N}}(Z^{(i)} - \hat{\mu}_Z)(Y^{(i)} - \mathbb{E}[Y]), \qquad \hat{\mu}_Z = \frac{1}{\bar{N}}\sum_{i=1}^{\bar{N}} Z^{(i)}$$

3 Generate N iid replicas $(Z^{(i)}, Y^{(i)})$, $i = 1, \dots, N$ of (Z, Y)

4 Compute $\hat{\mu}_{\text{CV}} = \frac{1}{N}\sum_{i=1}^{N}(Z^{(i)} - \hat{\alpha}_{\text{opt}}(Y^{(i)} - \mathbb{E}[Y]))$

5 Output $\hat{\mu}_{\text{CV}}$ and a confidence interval based on $\hat{\sigma}_{CV}$.

The estimator $\hat{\mu}_{CV}$ is

- unbiased: $\mathbb{E}[\hat{\mu}_{CV}] = \mathbb{E}[\mathbb{E}[\hat{\mu}_{CV} \mid \hat{\alpha}_{opt}]] = \mu$
- with variance $\mathbb{V}\mathrm{ar}[\hat{\mu}_{\text{CV}}] = \frac{1}{N}\left(\mathbb{V}\mathrm{ar}\left[Z_{\alpha_{\text{opt}}}\right] + \mathbb{V}\mathrm{ar}[\hat{\alpha}_{\text{opt}}]\,\sigma^2_Y\right)$

Control variates – one shot algorithm

Alternatively to the pilot run, one may consider a one-shot strategy

Algorithm: Control variate – one shot

1 Generate N iid replicas $(Z^{(i)}, Y^{(i)})$, $i = 1, \dots, N$ of (Z, Y)

2 Estimate $\hat{\alpha}_{\text{opt}} = \frac{\hat{\sigma}^2_{ZY}}{\sigma^2_Y}$, with

$$\hat{\sigma}^2_{ZY} = \frac{1}{N-1}\sum_{i=1}^{N}(Z^{(i)} - \hat{\mu}_Z)(Y^{(i)} - \mathbb{E}[Y]), \qquad \hat{\mu}_Z = \frac{1}{N}\sum_{i=1}^{N} Z^{(i)}$$

3 Estimate $\hat{\mu}_{\text{CV}} = \frac{1}{N}\sum_{i=1}^{N}(Z^{(i)} - \hat{\alpha}_{\text{opt}}(Y^{(i)} - \mathbb{E}[Y]))$

4 Output $\hat{\mu}_{\text{CV}}$ and a confidence interval based on $\hat{\sigma}_{CV}$.

- This estimator is biased in general
- A CLT holds $\quad \sqrt{N}\frac{\hat{\mu}_{\text{CV}} - \mu}{\hat{\sigma}^2(Z_{\alpha_{\text{opt}}})} \xrightarrow{\text{d}} N(0,1)$ as $N \to \infty$ (exercise)

Example – option pricing

- S_t: value of an asset at time t, modeled by

$$dS_t = rS_t\,dt + \sigma S_t\,dW_t, \quad t \in (0, T]$$

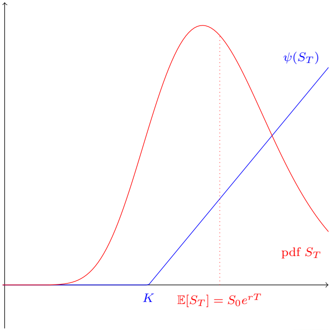
- call option: payoff $\psi(S_T) = (S_T - K)_+$:
- **Goal**: estimate (discounted) option price $\mu = \mathbb{E}[Z] = \mathbb{E}\left[e^{-rT}\psi(S_T)\right]$

Idea: use asset price S_T as control variate, with $\mathbb{E}[S_T] = S_0 e^{rT}$.

Control variate estimator

$$\hat{\mu}_{CV} = \frac{1}{N}\sum_{i=1}^{N} Z^{(i)} - \alpha(S_T^{(i)} - S_0 e^{rT})$$
$$= \hat{\mu}_Z - \alpha(\hat{\mu}_{S_T} - S_0 e^r T)$$

Z and S_T are positively correlated. If $\hat{\mu}_{S_T}$ is above the mean, then most likely also $\hat{\mu}_Z$ is above the mean and we have to correct with a negative term $-\alpha(\hat{\mu}_{S_T} - S_0 e^r T)$, with $\alpha > 0$.



Multiple control variates

In the control variate technique, one can actually use several control variates: $\boldsymbol{Y} = (Y_1, \ldots, Y_p)$ with known means $\mathbb{E}[Y_1], \ldots, \mathbb{E}[Y_p]$.

$$Z_\alpha = Z - \sum_{j=1}^{p} \alpha_j (Y_j - \mathbb{E}[Y_j]) = Z - \boldsymbol{\alpha} \cdot (\boldsymbol{Y} - \mathbb{E}[\boldsymbol{Y}])$$

Then

$$\begin{aligned} \mathbb{V}\mathrm{ar}[Z_\alpha] &= \mathbb{E}\left[(Z - \mu - \boldsymbol{\alpha} \cdot (\boldsymbol{Y} - \mathbb{E}[\boldsymbol{Y}]))^2\right] \\ &= \mathbb{V}\mathrm{ar}[Z] - 2 \underbrace{\mathrm{Cov}(Z, \boldsymbol{Y})}_{\in \mathbb{R}^p} \cdot \boldsymbol{\alpha} + \boldsymbol{\alpha}^\top \underbrace{\mathrm{Cov}(\boldsymbol{Y}, \boldsymbol{Y})}_{\in \mathbb{R}^{p \times p}} \boldsymbol{\alpha} \end{aligned}$$

Variance minimization:

$$\boldsymbol{\alpha}_{\mathrm{opt}} = \mathrm{Cov}(\boldsymbol{Y}, \boldsymbol{Y})^{-1} \mathrm{Cov}(Z, \boldsymbol{Y})$$

$$\mathbb{V}\mathrm{ar}\left[Z_{\alpha_{opt}}\right] = \mathbb{V}\mathrm{ar}[Z] - \mathrm{Cov}(Z, \boldsymbol{Y}) \, \mathrm{Cov}(\boldsymbol{Y}, \boldsymbol{Y})^{-1} \mathrm{Cov}(Z, \boldsymbol{Y})$$

$\boldsymbol{\alpha}_{opt}$ can be seen as the solution of a linear regression problem $Z - \mu \approx \boldsymbol{\alpha} \cdot (\boldsymbol{Y} - \mathbb{E}[Y])$ and Monte Carlo is done only on the residual $Z_a = Z - \boldsymbol{\alpha} \cdot (\boldsymbol{Y} - \mathbb{E}[Y])$.

Stratification

Standard setting

- $\boldsymbol{X}$ random vector in $\mathbb{R}^d$ with (joint) pdf $f : \Omega \subset \mathbb{R}^d \to \mathbb{R}_+$
- $Z = \psi(\boldsymbol{X}) \in \mathbb{R}$ output quantity
- **Goal**: compute $\mu = \mathbb{E}\left[Z\right] = \int \psi(\boldsymbol{x}) f(\boldsymbol{x}) d\boldsymbol{x}$

Idea: partition the sample space Ω in s non-overlapping strata $\Omega_1, \ldots, \Omega_s$ with $\mathbb{P}\left(X \in \Omega_j\right) = p_j$ known

Let

- $f_j(x) = \frac{1}{p_j} f(x) \mathbb{1}_{\Omega_j}(x)$: conditional density to $X \in \Omega_j$
- $X_j \sim f_j$: random variable taking values only in Ω_j
- $Z_j = \psi(X_j)$: output conditional on $X \in \Omega_j$

Then

$$\mu = \mathbb{E}\left[Z\right] = \sum_{j=1}^{s} \mathbb{E}\left[Z \mid X \in \Omega_j\right] \mathbb{P}\left(X \in \Omega_j\right) = \sum_{j=1}^{s} p_j \mathbb{E}\left[Z_j\right]$$

Idea: use independent Monte Carlo estimators for each $\mathbb{E}\left[Z_j\right]$

Stratification

Stratified estimator

$$\hat{\mu}_{\mathsf{Str}} = \sum_{j=1}^{s} p_j \hat{\mu}_j, \quad \hat{\mu}_j = \frac{1}{N_j} \sum_{i=1}^{N_j} Z_j^{(i)}, \qquad \text{with } Z_j^{(i)} \overset{\text{iid}}{\sim} Z_j$$

Properties

- $\hat{\mu}_{\mathsf{Str}}$ is unbiased. Indeed,

$$\mathbb{E}\left[\hat{\mu}_{\mathsf{Str}}\right] = \sum_{j=1}^{s} p_j \mathbb{E}\left[\hat{\mu}_j\right] = \sum_{j=1}^{s} p_j \mathbb{E}\left[Z_j\right] = \mathbb{E}\left[Z\right].$$

- $\mathbb{V}\mathrm{ar}\left[\hat{\mu}_{\mathsf{Str}}\right] = \sum_{j=1}^{s} p_j^2 \mathbb{V}\mathrm{ar}\left[\hat{\mu}_j\right] = \sum_{j=1}^{s} p_j^2 \frac{\mathbb{V}\mathrm{ar}[Z_j]}{N_j}$
- Let $N = \sum_{j=1}^{s} N_j$ and assume $\frac{N_j}{N} \to c_j > 0$ as $N \to \infty$. Then $\lim_{N\to\infty} N \mathbb{V}\mathrm{ar}\left[\hat{\mu}_{\mathsf{Str}}\right] = \sum_j p_j^2 \sigma_j^2 / c_j < +\infty$ and

$$\frac{\hat{\mu}_{\mathsf{Str}} - \mu}{\sqrt{\mathbb{V}\mathrm{ar}\left[\hat{\mu}_{\mathsf{Str}}\right]}} \xrightarrow{\text{d}} N(0,1), \quad \text{as } N \to \infty.$$

Computable $1-\alpha$ asymptotic confidence interval:

$$\hat{I}_\alpha = \left[\hat{\mu}_{\mathsf{Str}} - c_{1-\alpha/2}\hat{\sigma}_{\mathsf{Str}},\ \hat{\mu}_{\mathsf{Str}} + c_{1-\alpha/2}\hat{\sigma}_{\mathsf{Str}}\right]$$

Stratification

Algorithm: Stratification

```
for j = 1, ..., s do
    Generate N_j iid replicas Z_j^(i), i = 1, ..., N_j of Z_j
    Compute μ̂_j = (1/N_j) Σ_{i=1}^{N_j} Z_j^(i) and σ̂_j^2 = (1/(N_j−1)) Σ_{i=1}^{N_j} (Z_j^(i) − μ̂_j)^2
end
Compute μ̂_Str = Σ_{j=1}^{s} p_j μ̂_j and σ̂_Str^2 = Σ_{j=1}^{s} p_j^2 σ̂_j^2 / N_j
Output μ̂_Str and a confidence interval
```

for $j = 1, \ldots, s$ **do**

- Generate N_j iid replicas $Z_j^{(i)}$, $i = 1, \ldots, N_j$ of Z_j
- Compute $\hat{\mu}_j = \frac{1}{N_j} \sum_{i=1}^{N_j} Z_j^{(i)}$ and $\hat{\sigma}_j^2 = \frac{1}{N_j - 1} \sum_{i=1}^{N_j} (Z_j^{(i)} - \hat{\mu}_j)^2$

end

Compute $\hat{\mu}_{\mathrm{Str}} = \sum_{j=1}^{s} p_j \hat{\mu}_j$ and $\hat{\sigma}_{Str}^2 = \sum_{j=1}^{s} p_j^2 \frac{\hat{\sigma}_j^2}{N_j}$

Output $\hat{\mu}_{\mathrm{Str}}$ and a confidence interval

$$\hat{I}_\alpha = [\hat{\mu}_{\mathrm{Str}} - c_{1-\alpha/2} \hat{\sigma}_{\mathrm{Str}},\ \hat{\mu}_{\mathrm{Str}} + c_{1-\alpha/2} \hat{\sigma}_{\mathrm{Str}}]$$

Stratification – proportional allocation

Question: for a given budget $N = \sum_{j=1}^{s} N_j$, how to choose N_j?

Proportional allocation: $N_j = N p_j$

$$\Longrightarrow \qquad \mathbb{V}\mathrm{ar}\left[\hat{\mu}_{\mathsf{Str}}\right] = \sum_{j=1}^{s} p_j^2 \frac{\mathbb{V}\mathrm{ar}\left[Z_j\right]}{N_j} = \frac{1}{N} \sum_{j=1}^{s} p_j \mathbb{V}\mathrm{ar}\left[Z_j\right]$$

Interpretation: define $J \in \{1, \dots, s\}$ such that $J = j \iff \{X \in \Omega_j\}$.

$$\Longrightarrow \qquad \mathbb{V}\mathrm{ar}\left[\hat{\mu}_{\mathsf{Str}}\right] = \frac{1}{N} \sum_{j=1}^{s} p_j \mathbb{V}\mathrm{ar}\left[Z \mid J = j\right] = \frac{1}{N} \mathbb{E}_J[\mathbb{V}\mathrm{ar}\left[Z \mid J\right]]$$

Recal total variance formula $\mathbb{V}\mathrm{ar}\left[Z\right] = \mathbb{V}\mathrm{ar}\left[\mathbb{E}\left[Z \mid J\right]\right] + \mathbb{E}\left[\mathbb{V}\mathrm{ar}\left[Z \mid J\right]\right]$

$$\Longrightarrow \qquad \mathbb{V}\mathrm{ar}\left[\hat{\mu}_{\mathsf{Str}}\right] = \frac{1}{N}\left(\mathbb{V}\mathrm{ar}\left[Z\right] - \mathbb{V}\mathrm{ar}\left[\mathbb{E}\left[Z \mid J\right]\right]\right) \le \frac{\mathbb{V}\mathrm{ar}\left[Z\right]}{N} = \mathbb{V}\mathrm{ar}\left[\hat{\mu}_{\mathsf{CMC}}\right]$$

Stratification with proportional allocation always leads to variance reduction.

Example – 1d integration

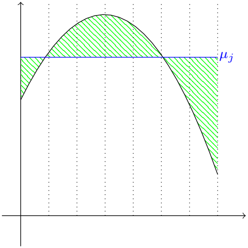
Let $X \sim \mathcal{U}(0,1)$, $\psi : [0,1] \to \mathbb{R}$, $Z = \psi(X)$. Goal: compute $\mu = \mathbb{E}[Z]$

Crude Monte Carlo

$$\hat{\mu}_{CMC} = \frac{1}{N} \sum_{i=1}^{N} \psi(X^{(i)}),$$

$$X^{(i)} \overset{iid}{\sim} \mathcal{U}(0,1)$$

$$\mathbb{V}\mathrm{ar}\left[\hat{\mu}_{CMC}\right] = \frac{\mathbb{V}\mathrm{ar}\left[Z\right]}{N}$$

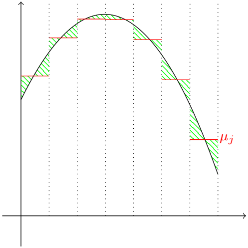


Stratification on s equal intervals

$$\hat{\mu}_{\mathrm{Str}} = \sum_{j=1}^{s} \sum_{i=1}^{N/s} \psi(X_j^{(i)}),$$

$$X_j^{(i)} \overset{\mathrm{iid}}{\sim} \mathcal{U}\left(\frac{j-1}{s}, \frac{j}{s}\right)$$

$$\mathbb{V}\mathrm{ar}\left[\hat{\mu}_{Str}\right] = \frac{1}{N} \sum_{j=1}^{s} \frac{\mathbb{V}\mathrm{ar}\left[Z_j\right]}{s}$$



EPFL

Example – 1d integration

Let $X \sim \mathcal{U}(0,1)$, $\psi : [0,1] \to \mathbb{R}$, $Z = \psi(X)$. Goal: compute $\mu = \mathbb{E}[Z]$

How many strata s and samples N_j per strata should we choose ?

Extreme cases:

- $s = 1, N_1 = N$ $\rightsquigarrow$ Crude Monte Carlo (No stratification)
- $s = N, N_j = 1$ $\rightsquigarrow$ randomized quadrature (mid point) rule.

It can be shown (exercise) that if $\psi \in C^1([0,1])$ and $s = N$ then

$$\mathrm{MSE}(\hat{\mu}_{Str}) = \mathbb{E}\left[(\mu - \hat{\mu}_{Str})^2\right] \leq \frac{1}{3N^3} \max_{x\in[0,1]} |\psi'(x)|^2$$

Convergence faster than CMC! but requires regularity of ψ

However, this result doesn't scale with the dimension. Consider $\psi : [0,1]^d \to \mathbb{R}$. If we stratify each variable with s strata we end up with s^d strata.

Assuming $\psi \in C^1([0,1]^d)$ and placing one point per stratum we have $\mathrm{MSE}(\hat{\mu}_{str}) \lesssim s^{-3} = N^{-\frac{3}{d}}$ ($\rightsquigarrow$ curse of dimensionality)

Example – stratification of Wiener process

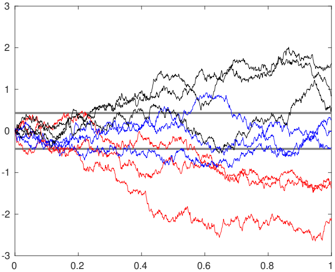
A more challenging stratification example: consider a Wiener process $\{W_t,\ t \in [0, T]\}$ and some function $Z = \phi(\{W_t\}_{t\in[0,T]})$, e.g. $Z = \max_{t\in[0,T]} W_t - \min_{t\in[0,T]} W_t$.

Goal: compute $\mu = \mathbb{E}[Z]$

Can we use stratification? How can we stratify the sample space of a Wiener process?

Idea:

- construct a stratified sample $\{W_T^{(i)}\}_{i=1}^N$ of $W_T \sim N(0, T)$.
- then, for each $W_T^{(i)}$ construct a Brownian Bridge $\{W_t^{(i)},\ t \in [0, 1] \mid W_T^{(i)}\}$



EPFL

Stratification – optimal allocation

Idea: find best choice of $\{N_j\}$ that minimizes variance of estimator

$$\{N_j^*\} = \underset{(N_1,\dots,N_s)}{\text{argmin}} \sum_{j=1}^{s} p_j^2 \frac{\mathbb{V}\mathrm{ar}\left[Z_j\right]}{N_j} \quad \text{such that} \quad \sum_{j=1}^{s} N_j = N.$$

Introducing Lagrangian function

$\mathcal{L}(\boldsymbol{N}, \lambda) = \sum_{j=1}^{s} p_j^2 \frac{\sigma_j^2}{N_j} + \lambda(\sum_{j=1}^{s} N_j - N)$ with $\sigma_j^2 = \mathbb{V}\mathrm{ar}\left[Z_j\right]$,

$$\frac{\partial \mathcal{L}}{\partial N_j} = -p_j^2 \frac{\mathbb{V}\mathrm{ar}\left[Z_j\right]}{N_j^2} + \lambda = 0 \quad \Longrightarrow \quad N_j \propto p_j \sigma_j$$

Optimal allocation: $N_j^* = \dfrac{N p_j \sigma_j}{\sum_{k=1}^{s} p_k \sigma_k}$

Optimal variance: $\mathbb{V}\mathrm{ar}\left[\hat{\mu}_{\mathrm{Str}}^*\right] = \frac{1}{N} \left(\sum_{j=1}^{s} p_j \sigma_j\right)^2$

Since optimal allocation is better than proportional allocation, it always achieves variance reduction !

Stratification – optimal allocation

In practice, a pilot run is needed to estimate the optimal allocation.

Algorithm: Stratification with optimal allocation

for $j = 1, \dots, s$ **do**
  Generate $\bar{N}_j$ iid replicas $Z_j^{(i)}, \quad i = 1, \dots, N_j$ of Z_j
  Estimate $\hat{\sigma}_j^2 = \frac{1}{\bar{N}_j - 1} \sum_{i=1}^{\bar{N}_j} (Z_j^{(i)} - \hat{\mu}_j)^2$
end
Choose $N = (c_{1-\alpha/2} \sum_{j=1}^{s} p_j \hat{\sigma}_j / \mathrm{tol})^2$ (to guarantee that $|\hat{I}_{\alpha,N}| < 2\mathrm{tol}$)
For $j = 1, \dots, s$, generate $N_j^* = \frac{N p_j \hat{\sigma}_j}{\sum_k p_k \hat{\sigma}_k}$ iid replicas $Z_j^{(i)}$ of Z_j
Compute $\hat{\mu}_i = \frac{1}{N_j^*} \sum_{i=1}^{N_j^*} Z_j^{(i)}$ and $\hat{\mu}_{\mathrm{Str}}^* = \sum_{j=1}^{s} p_j \hat{\mu}_j$

Latin Hypercube sampling

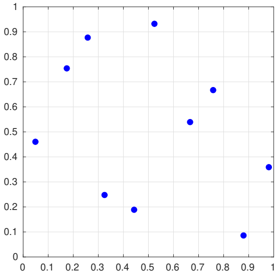
Consider

- $Z = \psi(X_1, \ldots, X_d)$ with $X_j \overset{\text{iid}}{\sim} \mathcal{U}([0,1])$
  (or more generally $X_j \sim f_j$ mutually independent)
- **Goal**: compute $\mathbb{E}[Z] = \int_{[0,1]^d} \psi(x_1, \ldots, x_d) dx_1 \cdots dx_d$

We could stratify each variable in s strata $\rightsquigarrow$ sample space $[0,1]^d$ stratified in s^d strata – unaffordable for d large.

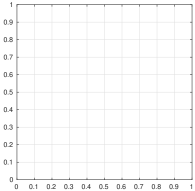
Idea: stratify the marginal distribution of each $X_j \sim \mathcal{U}([0,1])$ but not the joint distribution $X_j \sim \mathcal{U}([0,1]^2)$

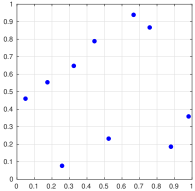
Latin Hypercube Sampling (LHS):
Generate a sample $\{\boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(N)}\}$ such that each variable X_j is stratified in N strata with one point per stratum.

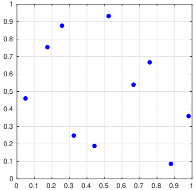


How to generate a LHS

- Divide hypercube in N^d blocks
- Place one uniformly distributed point in each diagonal block
- Permute columns
- Permute rows







LHS – Algorithm

Algorithm: LHS

1 Generate N iid points $\boldsymbol{U}^{(i)} \overset{\text{iid}}{\sim} \mathcal{U}((0,1)^d)$, $i = 1, \dots, N$

2 Generate d independent permutations π_j, $j = 1, \dots, d$ of $\{1, \dots, N\}$.
Let $\boldsymbol{\pi}^{(i)} = (\pi_1(i), \pi_2(i), \dots, \pi_d(i))$

3 Return $X^{(i)} = \frac{\boldsymbol{\pi}^{(i)} - 1 + \boldsymbol{U}^{(i)}}{N}$, $\quad i = 1, \dots, N$.

LHS estimator for $\mu = \mathbb{E}[Z] = \mathbb{E}[\psi(\boldsymbol{X})]$

$$\hat{\mu}_{LHS} = \frac{1}{N} \sum_{i=1}^{N} \psi(\boldsymbol{X}^{(i)})$$

LHS – Properties

- $X^{(i)} \sim \mathcal{U}((0,1)^d)$ (not independent, though)
- The LHS estimator is unbiased, $\mathbb{E}[\hat{\mu}_{\mathsf{LHS}}] = \mathbb{E}[\psi(X)]$.
- Result by [A. Owen 1997]

$$\mathbb{V}\mathrm{ar}[\hat{\mu}_{\mathsf{LHS}}] \leq \frac{\mathbb{V}\mathrm{ar}[Z]}{N-1}.$$

Hence $\lim_{N\to\infty} \mathbb{V}\mathrm{ar}[\hat{\mu}_{\mathsf{LHS}}] / \mathbb{V}\mathrm{ar}[\hat{\mu}_{\mathsf{CMC}}] \leq 1$.

EPFL

LHS – on variance of the estimator

- suppose ψ is of the form $\psi(\boldsymbol{X}) = \mu + \sum_{i=1}^{d} \psi_j(X_j)$
  Then, the LHS estimator corresponds to a stratified estimator with N strata on each function ψ_j $\quad\rightsquigarrow\quad$ super-canonical rate.

- For ψ arbitrary, define

$$\hat{\psi}_j(x_j) = \mathbb{E}\left[\psi(\boldsymbol{X}) - \mu \mid X_j = x_j\right] = \int_{[0,1]^{d-1}} (\psi(x_1, \ldots, x_d) - \mu)\, d\boldsymbol{x}_{\sim j}$$

with $\boldsymbol{x}_{\sim j} = (x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_d)$, and

$$\psi^{\mathrm{add}}(\boldsymbol{X}) = \mu + \sum_{i=1}^{d} \hat{\psi}_j(x_j).$$

Proposition. *[Stein 1987]*

$$\mathbb{V}\mathrm{ar}\left[\hat{\mu}_{LHS}\right] = \frac{\mathbb{V}\mathrm{ar}\left[\psi - \psi^{add}\right]}{N} + o\left(\frac{1}{N}\right).$$

Moreover, if $\mathbb{E}\left[\psi(\boldsymbol{X})^4\right] < +\infty$*, then*
$\sqrt{N}(\hat{\mu}_{LHS} - \mu) \xrightarrow{d} N(0, \mathbb{V}\mathrm{ar}\left[\psi - \psi^{add}\right])$ *as* $N \to \infty$.

EPFL

LHS – error estimation

The previous result can not be used to construct computable asymptotic confidence intervals as ψ^{add} is not know.

Simple idea to estimate the LHS error:

- Generate K independent LHS estimators $\hat{\mu}_{LHS}^{(i)},\ i = 1, \dots, K$
- compute the sample mean $\hat{\mu}_{\text{LHS}} = \frac{1}{K}\sum_{j=1}^{K} \hat{\mu}_{LHS}^{(j)}$
- estimate $\mathbb{V}\mathrm{ar}\left[\hat{\mu}_{LHS}\right]$ by sample variance estimator

---

Algorithm: LHS estimator

---

1. Generate K independent LHS designs $\{X^{(i,j)}\}_{i=1}^{N}$ of size N, for $j = 1, \dots, K$.
2. For each desing compute $\hat{\mu}_{LHS}^{(j)} = \frac{1}{N}\sum_{i=1}^{N} \psi(X^{(i,j)})$.
3. Compute $\hat{\mu}_{\text{LHS}} = \frac{1}{K}\sum_{j=1}^{K} \hat{\mu}_{LHS}^{(j)}$ and $\hat{\sigma}_{\text{LHS}}^2 = \frac{1}{K-1}\sum_{j=1}^{K} \left(\hat{\mu}_{LHS}^{(j)} - \hat{\mu}_{\text{LHS}}\right)^2$
4. Output $\hat{\mu}_{\text{LHS}}$ and the confidence interval $\hat{I}_\alpha = [\hat{\mu}_{\text{LHS}} - c_{1-\alpha/2}\frac{\hat{\sigma}_{\text{LHS}}}{\sqrt{K}},\ \hat{\mu}_{\text{LHS}} + c_{1-\alpha/2}\frac{\hat{\sigma}_{\text{LHS}}}{\sqrt{K}}]$.
