

MATH-414 – Stochastic simulation

Lecture 7: Variance Reduction Techniques III

Prof. Fabio Nobile

Outline

Control variates

Stratification

Latin Hypercube sampling

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

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

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

Do we get variance reduction?

Control variates - variance minimization

$$\text{Var} [\hat{\mu}_{CV}] = \frac{\text{Var}[Z_\alpha]}{N} \text{ and}$$

$$\text{Var} [Z_\alpha] = \text{Var} [Z] + \alpha^2 \text{Var} [Y] - 2\alpha \text{Cov}(Z, Y).$$

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

$$\frac{d\text{Var} [Z_\alpha]}{d\alpha} = 2\alpha \text{Var} [Y] - 2 \text{Cov}(Z, Y) = 0 \quad \implies \quad \alpha_{opt} = \frac{\text{Cov}(Z, Y)}{\text{Var} [Y]}$$

Minimal variance achievable

$$\text{Var} [Z_{\alpha_{opt}}] = \text{Var} [Z] - \frac{\text{Cov}(Z, Y)^2}{\text{Var} [Y]} = \text{Var} [Z] (1 - \rho_{ZY}^2)$$

with $\rho_{ZY}^2 = \frac{\text{Cov}(Z, Y)^2}{\text{Var}[Z]\text{Var}[Y]}$ correlation between Z and Y .

$\implies \text{Var} [Z_{\alpha_{opt}}] \leq \text{Var} [Z]$ and variance reduction increases as $\rho_{ZY} \rightarrow \{-1, 1\}$. **Ideal case:** $Y = \pm\gamma Z$ for which $\rho_{ZY} = \pm 1$ and $\text{Var} [Z_{\alpha_{opt}}] = 0!$

Useless since $\mathbb{E}[Y] = \gamma\mathbb{E}[Z]$ 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 $\hat{\alpha}_{opt} = \frac{\hat{\sigma}_{ZY}^2}{\sigma_Y^2}$ if σ_Y^2 known, or $\hat{\alpha}_{opt} = \frac{\hat{\sigma}_{ZY}^2}{\hat{\sigma}_Y^2}$ otherwise, with

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

- 3 Generate N iid replicas $(Z^{(i)}, Y^{(i)})$, $i = 1, \dots, N$ of (Z, Y)
 - 4 Compute $\hat{\mu}_{CV} = \frac{1}{N} \sum_{i=1}^N (Z^{(i)} - \hat{\alpha}_{opt}(Y^{(i)} - \mathbb{E}[Y]))$
 - 5 Output $\hat{\mu}_{CV}$ and a confidence interval based on $\hat{\sigma}_{CV}$.
-

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

- ▶ unbiased: $\mathbb{E}[\hat{\mu}_{CV}] = \mathbb{E}[\mathbb{E}[\hat{\mu}_{CV} \mid \hat{\alpha}_{opt}]] = \mu$
- ▶ if σ_Y known, variance $\text{Var}[\hat{\mu}_{CV}] = \frac{1}{N} (\text{Var}[Z_{\alpha_{opt}}] + \text{Var}[\hat{\alpha}_{opt}] \sigma_Y^2)$

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

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

- 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 $\hat{\sigma}_{\text{CV}}$.
-

- ▶ This estimator is **biased** in general
- ▶ A CLT holds $\sqrt{N} \frac{\hat{\mu}_{\text{CV}} - \mu}{\hat{\sigma}_{\text{CV}}} \xrightarrow{d} N(0, 1)$ as $N \rightarrow \infty$ (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

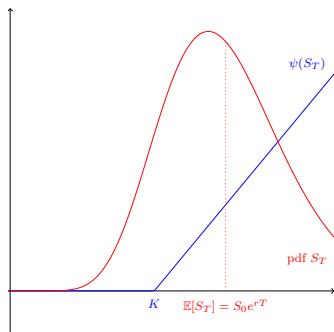
$$\mu = \mathbb{E}[Z] = \mathbb{E}[e^{-rT} \psi(S_T)]$$

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

Control variate estimator

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

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 $-\alpha(\hat{\mu}_{S_T} - S_0 e^{rT})$, with $\alpha > 0$.



Multiple control variates

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

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

Then

$$\begin{aligned} \text{Var}[Z_\alpha] &= \mathbb{E}[(Z - \mu - \alpha \cdot (\mathbf{Y} - \mathbb{E}[\mathbf{Y}]))^2] \\ &= \text{Var}[Z] - 2 \underbrace{\text{Cov}(Z, \mathbf{Y})}_{\in \mathbb{R}^p} \cdot \alpha + \alpha^\top \underbrace{\text{Cov}(\mathbf{Y}, \mathbf{Y})}_{\in \mathbb{R}^{p \times p}} \alpha \end{aligned}$$

Variance minimization:

$$\begin{aligned} \alpha_{\text{opt}} &= \text{Cov}(\mathbf{Y}, \mathbf{Y})^{-1} \text{Cov}(Z, \mathbf{Y}) \\ \text{Var}[Z_{\alpha_{\text{opt}}}] &= \text{Var}[Z] - \text{Cov}(Z, \mathbf{Y}) \text{Cov}(\mathbf{Y}, \mathbf{Y})^{-1} \text{Cov}(Z, \mathbf{Y}) \end{aligned}$$

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

Stratification

Standard setting

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

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

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

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

Stratification

Stratified estimator

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

Properties

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

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

- ▶ $\text{Var}[\hat{\mu}_{\text{Str}}] = \sum_{j=1}^s p_j^2 \text{Var}[\hat{\mu}_j] = \sum_{j=1}^s p_j^2 \frac{\text{Var}[Z_j]}{N_j}$
- ▶ Let $N = \sum_{j=1}^s N_j$ and assume $\frac{N_j}{N} \rightarrow c_j > 0$ as $N \rightarrow \infty$. Then $\lim_{N \rightarrow \infty} N \text{Var}[\hat{\mu}_{\text{Str}}] = \sum_j p_j^2 \sigma_j^2 / c_j < +\infty$ and

$$\frac{\hat{\mu}_{\text{Str}} - \mu}{\sqrt{\text{Var}[\hat{\mu}_{\text{Str}}]}} \xrightarrow{d} N(0, 1), \quad \text{as } N \rightarrow \infty.$$

Computable $1 - \alpha$ asymptotic confidence interval:

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

Stratification

Algorithm: Stratification

- 1 **for** $j = 1, \dots, s$ **do**
- 2 Generate N_j iid replicas $Z_j^{(i)}$, $i = 1, \dots, N_j$ of Z_j
- 3 Compute $\hat{\mu}_j = \frac{1}{N_j} \sum_{i=1}^{N_j} Z_j^{(i)}$ and $\hat{\sigma}_j^2 = \frac{1}{N_j-1} \sum_{i=1}^{N_j} (Z_j^{(i)} - \hat{\mu}_j)^2$
- 4 **end**
- 5 Compute $\hat{\mu}_{\text{Str}} = \sum_{j=1}^s p_j \hat{\mu}_j$ and $\hat{\sigma}_{\text{Str}}^2 = \sum_{j=1}^s p_j^2 \frac{\hat{\sigma}_j^2}{N_j}$
- 6 Output $\hat{\mu}_{\text{Str}}$ and a confidence interval

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

Stratification – proportional allocation

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

Proportional allocation: $N_j = N p_j$

$$\implies \text{Var} [\hat{\mu}_{\text{Str}}] = \sum_{j=1}^s p_j^2 \frac{\text{Var} [Z_j]}{N_j} = \frac{1}{N} \sum_{j=1}^s p_j \text{Var} [Z_j]$$

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

$$\implies \text{Var} [\hat{\mu}_{\text{Str}}] = \frac{1}{N} \sum_{j=1}^s p_j \text{Var} [Z \mid J = j] = \frac{1}{N} \mathbb{E}_J [\text{Var} [Z \mid J]]$$

Recal total variance formula $\text{Var} [Z] = \text{Var} [\mathbb{E} [Z \mid J]] + \mathbb{E} [\text{Var} [Z \mid J]]$

$$\implies \text{Var} [\hat{\mu}_{\text{Str}}] = \frac{1}{N} (\text{Var} [Z] - \text{Var} [\mathbb{E} [Z \mid J]]) \leq \frac{\text{Var} [Z]}{N} = \text{Var} [\hat{\mu}_{\text{CMC}}]$$

Stratification with proportional allocation always leads to variance reduction.

Example – 1d integration

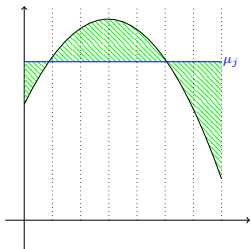
Let $X \sim \mathcal{U}(0, 1)$, $\psi : [0, 1] \rightarrow \mathbb{R}$, $Z = \psi(X)$. **Goal:** compute $\mu = \mathbb{E}[Z]$

Crude Monte Carlo

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

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

$$\text{Var}[\hat{\mu}_{CMC}] = \frac{\text{Var}[Z]}{N}$$

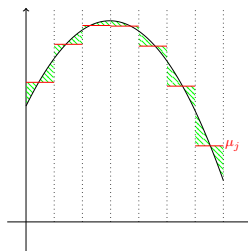


Stratification on s equal intervals

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

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

$$\text{Var}[\hat{\mu}_{Str}] = \frac{1}{N} \sum_{j=1}^s \frac{\text{Var}[Z_j]}{s}$$



Example – 1d integration

Let $X \sim \mathcal{U}(0, 1)$, $\psi : [0, 1] \rightarrow \mathbb{R}$, $Z = \psi(X)$. **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

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

Convergence faster than CMC! but requires regularity of ψ

However, this result doesn't scale with the dimension. Consider $\psi : [0, 1]^d \rightarrow \mathbb{R}$. 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 $\text{MSE}(\hat{\mu}_{str}) \lesssim s^{-3} = N^{-\frac{3}{d}}$ (\rightsquigarrow curse of dimensionality)

Example – stratification of Wiener process

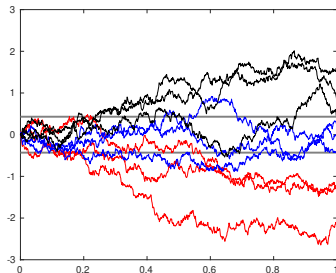
A more challenging stratification example: consider a Wiener process $\{W_t, t \in [0, T]\}$ and some function $Z = \phi(\{W_t\}_{t \in [0, T]})$, e.g.
 $Z = \max_{t \in [0, T]} W_t - \min_{t \in [0, T]} W_t$.

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 $\{W_t^{(i)}, t \in [0, 1] \mid W_T^{(i)}\}$



Stratification – optimal allocation

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

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

Introducing Lagrangian function

$$\mathcal{L}(\mathbf{N}, \lambda) = \sum_{j=1}^s p_j^2 \frac{\sigma_j^2}{N_j} + \lambda(\sum_{j=1}^s N_j - N) \quad \text{with} \quad \sigma_j^2 = \operatorname{Var}[Z_j],$$

$$\frac{\partial \mathcal{L}}{\partial N_j} = -p_j^2 \frac{\operatorname{Var}[Z_j]}{N_j^2} + \lambda = 0 \quad \implies \quad N_j \propto p_j \sigma_j$$

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

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

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

- 1 **for** $j = 1, \dots, s$ **do**
 - 2 Generate \bar{N}_j iid replicas $Z_j^{(i)}$, $i = 1, \dots, \bar{N}_j$ of Z_j
 - 3 Estimate $\hat{\sigma}_j^2 = \frac{1}{\bar{N}_j - 1} \sum_{i=1}^{\bar{N}_j} (Z_j^{(i)} - \hat{\mu}_j)^2$
 - 4 **end**
 - 5 Choose $N = (c_{1-\alpha/2} \sum_{j=1}^s p_j \hat{\sigma}_j / \text{tol})^2$ (to guarantee that $|\hat{I}_{\alpha, N}| < 2\text{tol}$)
 - 6 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
 - 7 Compute $\hat{\mu}_i = \frac{1}{N_j^*} \sum_{i=1}^{N_j^*} Z_j^{(i)}$ and $\hat{\mu}_{\text{Str}}^* = \sum_{j=1}^s p_j \hat{\mu}_j$
-

Latin Hypercube sampling

Consider

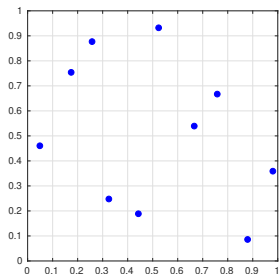
- ▶ $Z = \psi(X_1, \dots, X_d)$ with $X_j \stackrel{\text{iid}}{\sim} \mathcal{U}([0, 1])$
(or more generally $X_j \sim f_j$ mutually independent)
- ▶ **Goal:** compute $\mathbb{E}[Z] = \int_{[0,1]^d} \psi(x_1, \dots, x_d) dx_1 \cdots dx_d$

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 $X_j \sim \mathcal{U}([0, 1]^2)$

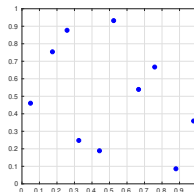
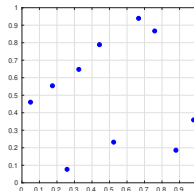
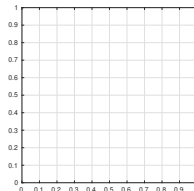
Latin Hypercube Sampling (LHS):

Generate a sample $\{\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(N)}\}$
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 $\mathbf{U}^{(i)} \stackrel{\text{iid}}{\sim} \mathcal{U}((0, 1)^d)$, $i = 1, \dots, N$
 - 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 $\mathbf{X}^{(i)} = \frac{\boldsymbol{\pi}^{(i)} - \mathbf{1} + \mathbf{U}^{(i)}}{N}$, $i = 1, \dots, N$.
-

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

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

LHS – Properties

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

$$\text{Var} [\hat{\mu}_{\text{LHS}}] \leq \frac{\text{Var} [Z]}{N - 1}.$$

Hence $\lim_{N \rightarrow \infty} \text{Var} [\hat{\mu}_{\text{LHS}}] / \text{Var} [\hat{\mu}_{\text{CMC}}] \leq 1$.

LHS – on the variance of the estimator

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

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

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

$$\psi^{\text{add}}(\mathbf{X}) = \mu + \sum_{i=1}^d \hat{\psi}_i(x_i).$$

Proposition. [Stein 1987, Loh 1996]

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

Moreover, if $\mathbb{E}[|\psi(\mathbf{X})|^3] < +\infty$, then

$\sqrt{N}(\hat{\mu}_{LHS} - \mu) \xrightarrow{d} N(0, \text{Var}[\psi - \psi^{\text{add}}])$ as $N \rightarrow \infty$.

LHS – error estimation

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

Simple idea to estimate the LHS error:

- ▶ Generate K independent LHS estimators $\hat{\mu}_{LHS}^{(i)}$, $i = 1, \dots, K$
- ▶ compute the sample mean $\hat{\mu}_{LHS} = \frac{1}{K} \sum_{j=1}^K \hat{\mu}_{LHS}^{(j)}$
- ▶ estimate $\text{Var}[\hat{\mu}_{LHS}]$ 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 design compute $\hat{\mu}_{LHS}^{(j)} = \frac{1}{N} \sum_{i=1}^N \psi(X^{(i,j)})$.
- 3 Compute $\hat{\mu}_{LHS} = \frac{1}{K} \sum_{j=1}^K \hat{\mu}_{LHS}^{(j)}$ and
$$\hat{\sigma}_{LHS}^2 = \frac{1}{K-1} \sum_{j=1}^K \left(\hat{\mu}_{LHS}^{(j)} - \hat{\mu}_{LHS} \right)^2$$
- 4 Output $\hat{\mu}_{LHS}$ and the confidence interval
$$\hat{I}_\alpha = \left[\hat{\mu}_{LHS} - c_{1-\alpha/2} \frac{\hat{\sigma}_{LHS}}{\sqrt{K}}, \hat{\mu}_{LHS} + c_{1-\alpha/2} \frac{\hat{\sigma}_{LHS}}{\sqrt{K}} \right].$$

One can replace \hat{I}_α with a CI based on the quantiles of the Student's t distribution, thanks to the asymptotic normality of each $\hat{\mu}_{LHS}^{(j)}$.