Übersicht

Conditional Monte Carlo

- Main Idea: Instead of computing $\theta = E[h(Y)] = E[X]$, choose a r.v. Z and:
  - Set $V := E[X|Z] = g(Z)$ for some function g.
  - Estimate $\theta := E[V]$.
- Conditions required:
  - Z can be easily simulated.
  - V=$g(Z)$ can be easily computed. For example: available in a closed form.

Variance reduction

- Assume $X \in L^2(\Omega)$, then
  $X - E[X] - (E[X|Z] - E[X]) \perp E[X|Z] - E[X]$,
  we have the classical decomposition of the variance

$$Var(X) = E[Var(X|Z)] + Var(E[X|Z]),$$

  where $Var(X|Z) := E\left[(X - E[X|Z])^2 \middle| Z\right]$ is a non-negative r.v.
  Thus

$$Var(X) \geq Var(E[X|Z]).$$

- Note that the r.v. X and Z should be dependent in order to achieve a variance reduction.

Example

- Aim: Estimate $\theta := P(U + Z > 4)$, $U \sim Exp(1)$ and $Z \sim Exp(1/2)$.
  Setting $X := 1_{\{U+Z>4\}}$, then $\theta := \mathbb{P}(U + Z > 4) = E[X]$.
- Classical Monte Carlo method:
  - Generate $U_1, \ldots, U_n$ and $Z_1, \ldots, Z_n$ independently.
  - Set for all i, $X_k := 1_{\{U_k+Z_k>4\}}$
  - $\hat{\theta}_n := \frac{1}{n}\sum_{k=1}^{n} X_k$.
  - Confidence intervals are computed as usual.

Example

- Set $V := E[X|Z] = g(Z)$. Then by direct computation,

  $$E[X|Z=z] = 1-F_U(4-z) = \exp-(4-z)1_{\{0\leq z\leq 4\}}+1_{\{z>4\}},$$

  where F_U is the cumulative distributin function (CDF) of U.
- Conditional Monte Carlo method for $E[V]$:
  - Generate $Z_1, \dots, Z_n$ independently.
  - Set for all i, $V_k := E[X|Z_k] = g(Z_k)$.
  - $\hat{\theta}_{n,CMC} := \frac{1}{n}\sum_{k=1}^n V_k$.
  - Confidence intervals are computed as usual, but using $(V_k)_k$ instead of $(X_k)_k$.

Comments

- Conditional Monte Carlo works well when we can compute exactly the function $g(.)$. This is not always possible.
- An alternative could be to combine some other variance reduction methods with Conditional Monte Carlo to estimate $\hat{\theta}_{n,CMC}$. For example, antithetic variables when $g(.)$ is monotonic. Etc...

Übersicht

Stratified Sampling

- Main Idea: use conditioning to reduce the variance.
- Tools: Law of total expectation, Conditional sampling.
- Estimation for $E[h(Y)]$, using conditional sampling w.r.t. a second r.v. W.
  Without loss of generality, we will treat the estimation of $E[Y]$.

Stratified Sampling: Motivation

- Consider the following 2-step random experience in :
  We have an urn with four different colored balls: red, green, blue and orange (the colors will be referred to as r, g, b and o):

Step 1) Pick at random a colored ball from the urn. The picked color will be referred to by c. We denote by I the r.v. describing this Step.

Step 2) Conditional on the color $I = c$ having been drawn in Step 1), we receive a payoff Y which is drawn from the pdf $f_c(.)$.

- How much would we win on average?

Stratified Sampling: Motivation

Before Answering this question, let's describe the law μ of the payoff Y.

- We denote by ν the discrete law of I and μ_i the law with pdf f_i. Then for all $A \in \mathcal{B}(\mathbb{R})$,

$$\begin{aligned} \mu(A) &= \sum_{i \in \mathcal{I}} \nu(\{i\}) \mu_i(A) \\ &= \sum_{i \in \mathcal{I}} \nu(\{i\}) \int_{\mathbb{R}} 1_{\{A\}} f_i(y) dy. \end{aligned}$$

Stratified Sampling: Motivation

- Estimation for the gain:
  2-Step Algorithm for n simulations of Y:

Step (1) Draw n .i.i.d. r.v. $(I_k)_{1\leq k\leq n}$ according to the discrete law ν. These r.v. are valued in $\mathcal{I} := \{\mathrm{r,g,b,o}\}$.

Step (2) For each $k = 1, \ldots, n$, given $I_k = i$, simulate independently the r.v. Y_k from $f_i(.)$.

- $\hat{\theta}_n := \frac{1}{n}\sum_{k=1}^{n} Y_k$.
- Assume that taking each one of the four colors has probability $\frac{1}{4}$. Assume in addition that $n = 1000$ and that the color r has been choosen 246 times, g 270 times, b 226 times and o 258 times. Finally, assume that f_g leads to high payoffs and f_b to low ones. Would this influence the confidence in $\hat{\theta}_n$?

Stratified Sampling: Motivation

$$
\begin{aligned}
E[Y] &= E[E[Y|I]] \\
&= \frac{1}{4}E[Y|I=r] + \frac{1}{4}E[Y|I=g] \\
&+ \frac{1}{4}E[Y|I=b] + \frac{1}{4}E[Y|I=o] \\
&= \frac{1}{4}\theta_r + \frac{1}{4}\theta_g + \frac{1}{4}\theta_b + \frac{1}{4}\theta_o.
\end{aligned}
$$

We can define the estimator

$$
\hat{\theta}_{st,n} = \frac{1}{4}\hat{\theta}_{st,n_r} + \frac{1}{4}\hat{\theta}_{st,n_g} + \frac{1}{4}\hat{\theta}_{st,n_b} + \frac{1}{4}\hat{\theta}_{st,n_o},
$$

whith $n_r + ng + n_b + n_o = n$.
Recall that $Var(Y) = E[Var(Y|I)] + Var(E[Y|I])$. Thus

$$
Var(Y) \geq E[Var(Y|I)]
$$

$$
\begin{aligned}
E[Var(Y|I)] &= \frac{1}{4}Var(Y|I=r) + \frac{1}{4}Var(Y|I=g) \\
&+ \frac{1}{4}Var(Y|I=b) + \frac{1}{4}Var(Y|I=o) \\
&= \frac{1}{4}Var(Y_r + Y_g + Y_b + Y_o).
\end{aligned}
$$

A fair comparison of $Var(\hat{\theta}_n)$ and $Var(\hat{\theta}_{st,n})$ should compare $Var(Y^1 + Y^2 + Y^3 + Y^4)$ and $Var(Y_r + Y_g + Y_b + Y_o)$, where Y^1, Y^2, Y^3 and Y^4 are four i.i.d. random variables drawn independently from the 2-Step algorithm. We get

$$
Var(Y^1 + Y^2 + Y^3 + Y^4) = 4Var(Y) \geq Var(Y_r + Y_g + Y_b + Y_o).
$$

Stratified Sampling

- Y is a random variable. The aim is to estimate $\theta = E[Y]$. Let W be a real r.v. and $(\Delta_i)_{1\le i\le m}$ be a finite partition of $\mathbb{R}$. Now, setting $I := \sum_{i=1}^m i 1_{\{W\in\Delta_i\}}$, we get

$$\begin{aligned} E[Y] &= E[E[Y|I]] = \sum_{i=1}^m P(I=i)E[Y|I=i] \\ &= \sum_{i=1}^m P(W\in\Delta_i)E[Y|W\in\Delta_i] \\ &= \sum_{i=1}^m p_i E_i. \end{aligned}$$

where $p_i := P(W\in\Delta_i)$ and $E_i := E[Y|W\in\Delta_i]$.

Stratified Sampling

- The random variable W is called *" the stratifying variable"* and the subsets $(\Delta_i)_{1 \leq i \leq m}$ are called *" the strata"*.
- In our example, the event $\{W \in \Delta_i\}$ equals $\{I = i\}$ and $m = |\mathcal{I}|$.

- Conditions required:
  - p_i can be easily computed.
  - It is easy to generate Y given $\{W \in \Delta_i\}$.

Stratified Sampling

- Consider m independent random variables $Y^{(i)} \sim \mathcal{L}(Y|W \in \Delta_i)$ the conditional law of Y given $\{W \in \Delta_i\}$ and $\theta_i := E[Y^{(i)}]$. Then

$$\theta = E[Y] = E[E[Y|I]] = \sum_{i=1}^{m} p_i E_i,$$
$$= \sum_{i=1}^{m} p_i \theta_i.$$

$\rightarrow$ In order to estimate θ, we need to estimate θ_i, using n_i independent samples of $Y^{(i)}$. The estimate of θ is given by:

$$\hat{\theta}_{st,n} = \sum_{i=1}^{m} p_i \hat{\theta}_{i,n_i}.$$

- $\hat{\theta}_{st,n}$ is unbiased if for all i, $\hat{\theta}_{i,n_i}$ is unbiased.

Variance reduction: Sub-optimal allocation approach

- Determine $(n_i)_i$, to achieve some variance reduction.
- Optimal approach: Take $(n_i)_i$ minimizing the variance $Var(\hat{\theta}_{st,n})$ subject to constraint $\sum_{i=1}^{m} n_i = n$, for given number n of overall samples.
- Sub-optimal allocation: take for all i, $n_i = np_i$.
  This last choice gives

$$Var(\hat{\theta}_{st,n}) \leq \frac{\sigma^2}{n} =: Var(\hat{\theta}_n),$$

where $\sigma := Var(Y)$ and $\hat{\theta}_n$ is the classical simulation estimator of θ.

Variance reduction: Optimal variance reduction

- $$Var(\hat{\theta}_{st,n}) = \sum_{i=1}^{m} \frac{p_i^2 \sigma_i^2}{n_i}, \text{ where } \sigma_i := Var(Y^{(i)}).$$

  $\rightarrow$ Constrained optimization problem:

  $$\min_{n_i} \{ \sum_{i=1}^{m} \frac{p_i^2 \sigma_i^2}{n_i} \} \text{ subject to } \sum_{i=1}^{m} n_i = n$$

- The solution is given by:

  $$n_i^* := \frac{p_i \sigma_i}{\sum_{i=1}^{m} p_i \sigma_i} n$$

  The minimal variance is

  $$Var(\hat{\theta}_{st,n^*}) = \frac{(\sum_{i=1}^{m} p_i \sigma_i)^2}{n}$$

Example 1: Simulation in the conditional laws

- Want to estimate $\theta := E[\sqrt{1-U^2}]$, where $U \sim U(0,1)$.
- Take $W = U$ as a stratification variable. This is a possible since
  - $P(W \in \Delta_i)$ can be easily computed.
  - The conditional law $\mathcal{L}(Y|W \in \Delta_i)$ can be easily generated.
- $\mathcal{L}(U|U \in [a,b]) = U(a,b)$.
  $\rightarrow$ By choosing N_{st} equiprobable strates, so that $\Delta_i = [\frac{i-1}{N_{st}}, \frac{i}{N_{st}}]$, for $i = 1, \dots, Nst$:
  - $P(W \in \Delta_i) = \frac{1}{N_{st}}$.
  - $U^{(i)} \sim \mathcal{L}(U|U \in \Delta_i)$. Then $U^{(i)} \sim U(\frac{i-1}{N_{st}}, \frac{i}{N_{st}})$

Example 1: Matlab code and experiments

- Matlab code
  $function[theta, var] = strat(Mtot, Nst)$
  $p = 1/(Nst);$
  $Mst = Mtot/Nst;$
  $theta = 0; var = 0;$
  $forj = 1 : Nst$
  $U = (j-1)/Nst + rand(Mst, 1)/Nst;$
  $X = sqrt(1 - U.^2);$
  $theta = theta + p * mean(X);$
  $Sum = sum(X);$
  $Sum_{squares} = sum(X2);$
  $Sig_{squares,j} = (Sum_{squares} - (Sum^2)/Mst)/(Mst - 1);$
  $var = var + (Sig_{squares,j} * p)^2/Mst;$
  $end;$
  $CI = [theta - 1.96 * sqrt(var), theta + 1.96 * sqrt(var)];$
  end

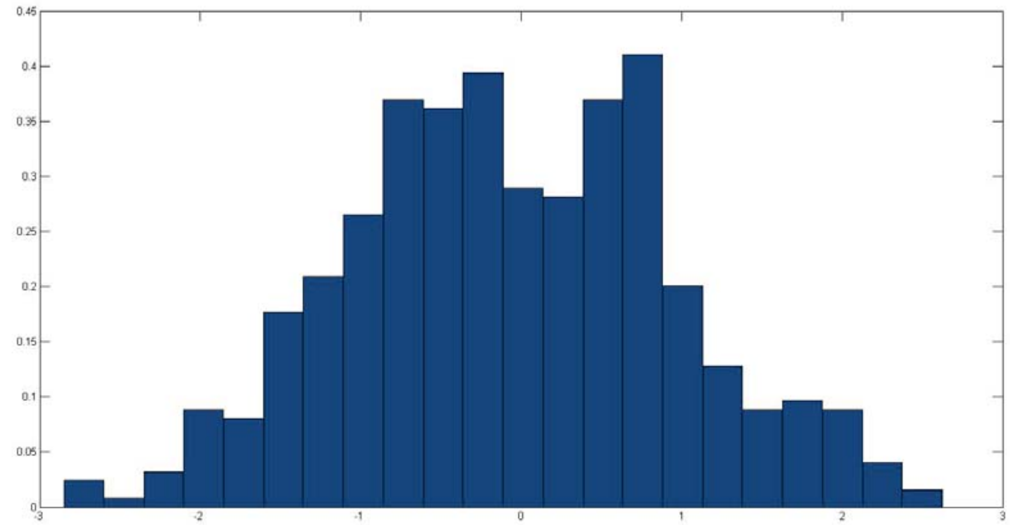
- Numerical experiments: For a sample of size 10^4
  10 strates: $\hat{\theta}_{st} = 0,784(1.2 \times 10^{-9})$.
  Without stratification: $\hat{\theta} = 0,7883(0.049)$.

Example 2: Stratifying a standard normal distribution

- For $X \sim N(0,1)$, how to simulate the conditional law $\mathcal{L}(X|X \in [a,b])$?
  $\rightarrow$ Let Φ be the CDF of X. If $\widetilde{U} \sim U(\Phi(a), \Phi(b))$, then $\widetilde{X} := \Phi^{-1}(\widetilde{U}) \sim \mathcal{L}(X|X \in [a,b])$.
- Exercise: Write a Matlab script giving the histogram for a stratified sample of X, using 100 equiprobable strates.

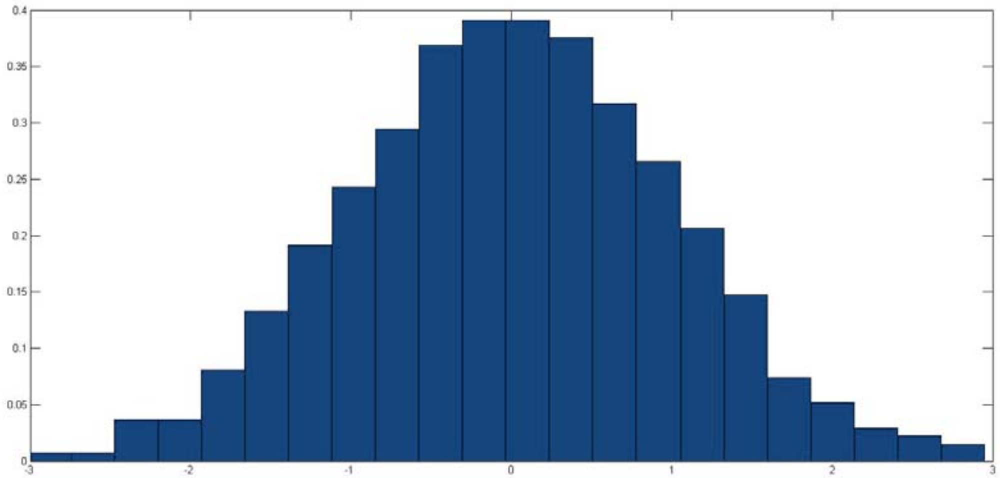
Example 2: Stratifying a standard normal distribution

Non stratified sample.



Example 2: Stratifying a standard normal distribution

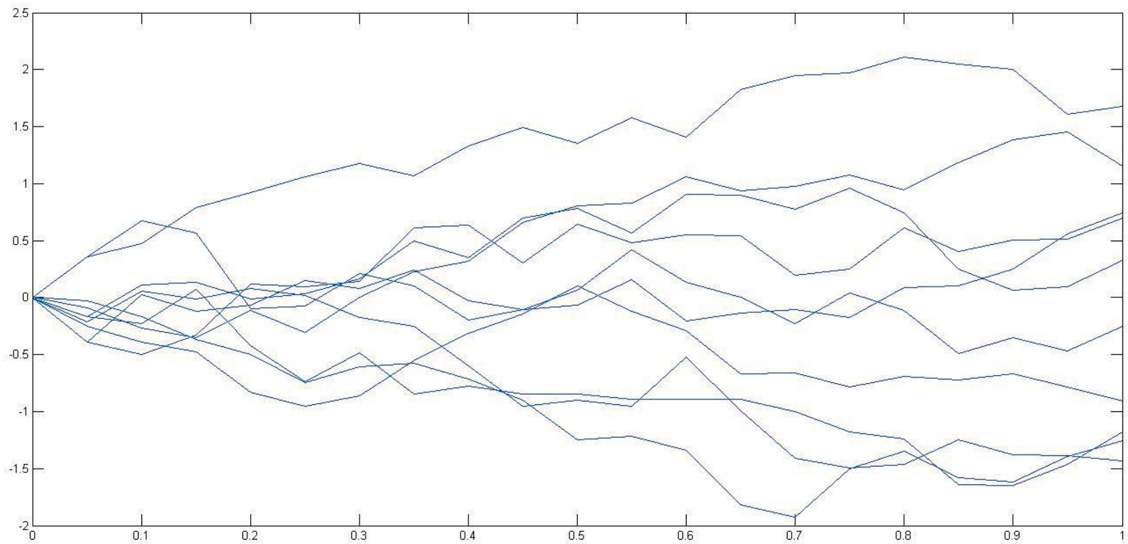
Stratified sample using 100 equiprobable strates.



Stratification of the terminal value of a BM

- Motivated by variance reduction in Option pricing.
- Based on the Levy construction of the Brownian Motion, after stratifying W_T.

Stratification of the terminal value of the Brownian motion.



Stratified sampling: Some comments

- Typically, $(\sigma_k)_k$ are unknown. Their numerical estimation has additional computational cost.
- In many cases, it is more convenient to use the sub-optimal variance reduction.

Literatur

- Sheldon Ross, Simulation, Academic Press, 2006
- Paul Glasserman, Monte Carlo Methods in Financial Engineering, Springer, 2004
- Higham+Higham, Matlab Guide