Reducing Variance via Stratified and Conditional Sampling

Middle East Technical University
Institute of Applied Mathematics

Fall Semester 2015

Introduction

Again it is the aim to estimate $\mu = \mathbb{E}(X)$.

This time we attempt to reduce the number of necessary Monte Carlo iterations by computing and exploiting conditional expectations $\mathbb{E}(X|Y)$, with X and Y being random variables.

Both following techniques are similar but make use of contrary assumptions.

Stratified Sampling Definition

- In stratified sampling, the population is partitioned into non-overlapping groups and these groups are called strata.
- A sample is selected by some design within each stratum.
- The most common strata used in stratified random sampling are age, gender, socioeconomic status, religion, nationality and educational attainment.
- Researchers generally use stratified random sampling in order to observe the relationships between two or more subgroups.
- In simple random sampling technique, we are not sure whether the sample represents equally or proportionally the subgroups.

Advantage and Disadvantage of Stratified Sampling

- In stratified sampling, it is possible to sample even the smallest and most inaccessible subgroups in the population.
- This situation allows us to sample the rare extremes of the given population.
- Stratified sample will always achieve greater precision than a simple random sample, if the strata have been chosen so that members of the same stratum are as similar as possible in terms of the characteristic of interest.
- The greater the differences between the strata, the greater the gain in precision.
- The main disadvantage of stratified random sampling is that it is difficult to identify appropriate strata.

Proportionate Sampling and Disproportiona Sampling

- The important thing in proportionate sampling is to use the same sampling fraction for each stratum regardless of the differences in population size of the strata.
- The only difference between proportionate and disproportionate stratified random sampling is their sampling fractions. With disproportionate sampling, the different strata have different sampling fractions.
- The purpose of disproportionate sampling can be explained by the following example.

Disproportional Stratified Sampling

- Suppose we are searching for the amount of average consumption of food and beverages.
- In that sense, we know the number of grocery stores and supermarkets. Assume 80% of the total stores are grocery and the rest is supermarket.
- However, we also know that most of the sales occur through supermarkets suppose 80%.
- If we use proportionate sampling, we tend to give higher weight to grocery store sales although in total sales its share is much lower compared to supermarkets.
- Therefore, we employ higher sampling friction to supermarkets, 80%, whereas 20% to grocery stores.

Application of Stratified Sampling in Finance

- Stratified sampling is based on dividing an index into manageable risk elements, such as a bond index.
- The multiple dimensions of risk within a bond portfolio are commonly defined as follows: currency, yield curve, duration, sector, credit, issuer and liquidity.
- Purpose is to create a portfolio which replicates the bond index.
- Stratified sampling is one of the most efficient fixed income index management techniques when full replication is not an option.
- Rather than holding all the thousands of bonds in the specific bond index, a portfolio manager creates a sample replication of the index.

Application of Stratified Sampling in Bond Index

The replication relies on the following steps:

- Screening all of the bonds in the universe using a broad set of analytics
- Finding out the dimensions (stratas) such as currency, liquidity, duration and credit risk for fixed income
- Use stratified sampling approach based on the stratas

The discrepancy between the bond index and replicating portfolio is called tracking error. Purpose for passive funds is to have low tracking error.

Deriving the Estimator

stratified MC estimator:

X is divided into d different parts (strata) determined by $y_1, ..., y_d$.
Assume all $p_i = P(Y = y_i)$ are known. Further denomination for $i = 1, ..., d$:

$$\overline{X}_{i,N_i} := \frac{1}{N_i} \sum_{j=1}^{N_i} X_j^{(i)} \ , \ \mu_i := \mathbb{E}(X|Y = y_i) \ , \ \sigma_i^2 := Var(X|Y = y_i).$$

Then with $N = \sum_{i=1}^{d} N_i$ and all $X_j^{(i)}$ having same distribution as $X|Y = y_i$:

$$\boxed{\overline{X}_{strat,N} = \sum_{i=1}^{d} p_i \overline{X}_{i,N_i}} \ .$$

Variance of stratified estimator

$$Var(\overline{X}_{strat,N}) = Var(\sum_{i=1}^{d} p_i \overline{X}_{i,N_i}) = \sum_{i=1}^{d} p_i^2 Var(\overline{X}_{i,N_i}) = \sum_{i=1}^{d} p_i^2 \frac{\sigma_i^2}{N_i} = \frac{\sigma^2(q)}{N}$$

with $\sigma^2(q) = \sum_{i=1}^{d} p_i^2 \frac{\sigma_i^2}{q_i}$ and $q_i = \frac{N_i}{N}$.

- In case of proportional stratification: $q_i = p_i$ and $Var(\overline{X}_{strat,N}) = \sum_{i=1}^{d} p_i \frac{\sigma_i^2}{N} \leq \frac{\sigma^2}{N} = Var(\overline{X}_{crude,N})$.
- $\sigma^2(q)$ can be optimised to determine stratification with highest variance reduction and results in: $q_i^* = \frac{p_i \sigma_i}{\sum_{k=1}^{d} p_k \sigma_k}$.

$$\Rightarrow Var(\overline{X}^*_{strat,N}) = \frac{\left(\sum_{i=1}^{d} p_i \sigma_i\right)^2}{N} \leq \sum_{i=1}^{d} \frac{p_i \sigma_i^2}{N}$$

Remark: Refining the strata yields in further variance reduction.

Confidence Interval - stratified sampling

Remembering properties of crude estimator:

$$\frac{1}{\sqrt{N_i}} \sum_{j=1}^{N_i} \left(X_j^{(i)} - \mu_i \right) \xrightarrow{D} \mathcal{N}(0, \sigma_i^2) \Rightarrow \overline{X}_{i,N_i} \sim \mathcal{N}\left(\mu_i, \frac{\sigma_i^2}{N_i} \right)$$

With all these subestimators being independent:

$$\overline{X}_{strat,N} \sim \mathcal{N}\left(\mu, \frac{1}{N} \sum_{i=1}^{d} \frac{p_i^2 \sigma_i^2 N}{N_i} \right)$$

and with $\frac{N_i}{N} \overset{N \to \infty}{\longrightarrow} p_i$:

$$\overline{X}_{strat,N} \sim \mathcal{N}\left(\mu, \frac{1}{N} \sum_{i=1}^{d} p_i \sigma_i^2 \right)$$

and by estimating strata variances σ_i^2 with:

$$\hat{\sigma}_i^2 = \frac{1}{N_i - 1} \sum_{j=1}^{N_i} \left(X_j^{(i)} - \overline{X}_{i,N_i} \right)^2$$

we get: $\left[\overline{X}_{strat,N} - \frac{z_{1-\alpha/2}}{\sqrt{N}} \sqrt{\sum_{i=1}^{d} \hat{\sigma}_i^2 p_i} \, , \, \overline{X}_{strat,N} + \frac{z_{1-\alpha/2}}{\sqrt{N}} \sqrt{\sum_{i=1}^{d} \hat{\sigma}_i^2 p_i} \right]$.

Example for solving higher dimensional Integral

$$\int_{[0,1]^{30}} x_1^2 + \cdots + x_{30}^2 \, \mathrm{d}\mu = 10$$

RNs	strata	crudeEst	stratEst	Var(crudeEst)	Var(stratEst)
10	10	9.0106	9.9717	10.4752	0.1379
100	10	9.2259	9.8620	1.3242	0.0146
	20		10.0075		0.0021
	100		9.9907		1.0234 * 10^{-4}
1000	10	9.9694	9.9990	0.0803	0.0013
	20		9.9927		1.7725 * 10^{-4}
	100		9.9991		7.9662 * 10^{-6}

20 replications are used for estimating variance.

The Method of Poststratification

The probabilities p_i of the different strata A_i are known, but a conditional sampling on the different strata is not being interested in. For example, in a telephone interview the respondents can not be placed into a male or female strata until after the respondent is contacted. Steps of method are as following:

- Generate a sample $X_1, \dots, X_N$ in the usual way.
- Classify the different values of this sample in groups related to the different strata A_i, i.e. assign each X_j to the sample A_i that is related to it. Set $N_i = |X_j \in A_j : j = 1, \dots, N|$.
- On each A_i calculate the crude Monte Carlo estimator with the assigned X values of the sample to obtain the strata means $\hat{X}_{i,N_i}$.
- Obtain the poststratified Monte Carlo estimator $\hat{X}_{strat,N}$ as

$$\hat{X}_{strat,N} = \sum_{i=1}^{d} p_i \hat{X}_{i,N_i},$$

with $p_i = \mathbb{P}(A_i)$.

Poststratification Example

Poststratification (stratification after the selection of a sample) is often appropriate when a simple random sample is not properly balanced by the representation.

Ex. Estimate the average weight and take a simple random sample of 100 people with following information:

Table: Related Information of Example

Male	**Female**
n_1=20	n_2=80
$\bar{y}_1 = \hat{X}_{1,N_1} = 180$ lbs.	$\bar{y}_2 = \hat{X}_{2,N_2} = 120$ lbs.

The overall sample mean is 132. This is obviously not balanced with respect to gender. This is likely an underestimate due to the under representation of males in the data. How can we account for this?

Poststratification Example (Cont.)

In the population $\frac{N_1}{N} = 0.5$ and $\frac{N_2}{N} = 0.5$. At that point note that the crude Monte Carlo estimator of the full sample is given as:

$$\bar{y}_{st} = \bar{X}_N = \frac{1}{N}\sum_{i=1}^{N} X_i = \frac{1}{N}\sum_{i=1}^{d} N_i \hat{X}_{i,N_i} = \sum_{i=1}^{d} \frac{N_i}{N}\hat{X}_{i,N_i}.$$

Therefore we have:

$$\bar{y}_{st} = 0.5 \cdot \bar{y}_1 + 0.5 \cdot \bar{y}_2 = \frac{N_1}{N}\bar{y}_1 + \frac{N_2}{N}\bar{y}_2 = 150.$$

Poststratification Example II

A firm knows that 40% of its accounts receivable are wholesale and 60% are retail. However, to identify an account without pulling a file and looking at it is difficult. An auditor randomly sampled 100 accounts without replacement. Here are the results of his sampling:

Table: Samplings of Auditor

Wholesale	**Retail**
n_1=70	n_2=30
$\bar{y}_1 = \hat{X}_{1,N_1} = 520$	$\bar{y}_2 = \hat{X}_{2,N_2} = 280$
$s_1 = 210$	$s_2 = 90$

Poststratification Example II (Cont.)

$$\bar{y}_{st} = \frac{N_1}{N}\bar{y}_1 + \frac{N_2}{N}\bar{y}_2 = 0.4 \cdot 520 + 0.6 \cdot 280 = 376,$$

and

$$Var(post - stratified(\bar{y})) \approx$$

$$\frac{1}{n}\left(\frac{N_1}{N}s_1^2 + \frac{N_2}{N}s_2^2\right) + \frac{1}{n^2}\left[\left(1 - \frac{N_1}{N}\right)s_1^2 + \left(1 - \frac{N_2}{N}\right)s_2^2\right]$$

$$\frac{1}{100}\left[0.4 \times (210)^2 + 0.6 \times (90)^2\right] + \frac{1}{100^2}[0.6 \times (210)^2 + 0.4 \times (90)^2] =$$

$$225 + 2.97 = 227.97$$

Latin Hypercube Sampling

Latin hypercube sampling is an extension of stratification for sampling in multiple dimensions.

- For each random variable, divide range into n non-overlapping intervals
- Select a random value in each interval for each variable
- Randomly select one value from each list and use in MC calculation
- Idea is eliminating some of variance between the strata by a specific way of multidimensional sampling

Latin Hypercube Sampling

For N random vectors in d-dimension, we have main steps as:

- Stratify each component $X^{(j)}, j = 1, \ldots, d$ with N equally likely strata $A_1, \ldots, A_N$.
- Sample exactly one observation $Y_i^{(j)}$ from stratum A_i for all $i = 1, \ldots, N$ for each component $j = 1, \ldots, d$.
- Choose d permutations $\pi_1, \ldots, \pi_d$ randomly from the set of permutations of $1, 2, \ldots, N$.
- Set $X_i^{(j)} = Y_{\pi_j(i)}^{(j)}$ for $i = 1, \ldots, N$, $j = 1, \ldots, d$.

Latin Hypercube Sampling

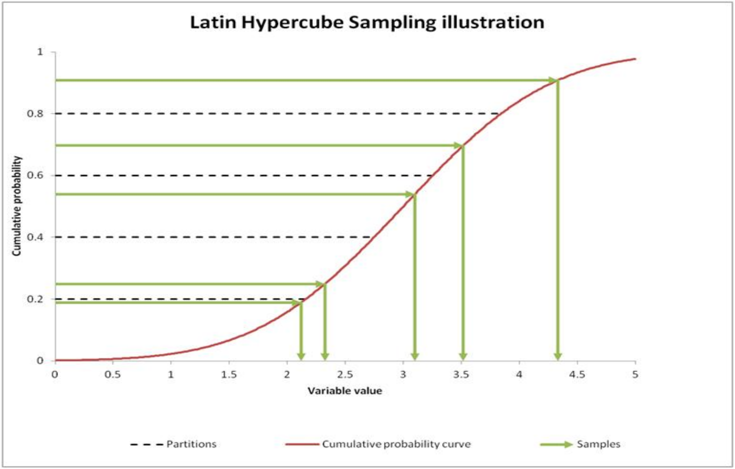


Figure: Latin Hypercube Sampling

Example of LHS

Simulation of 4 vectors $X_i \sim \mathcal{U}[0,1]^3$ by LHS:

- Each $X^{(j)}$ is stratified into

$$A_1 = [0, 0.25], A_2 = (0.25, 0.5], A_2 = (0.5, 0.75], A_4 = (0.75, 1].$$

- We then sample all $Y_i^{(j)}$ by stratified sampling to obtain:

$$Y^{(1)} = (0.095500046, 0.493293558, 0.701216163, 0.866725669)$$

$$Y^{(2)} = (0.025170141, 0.349131748, 0.705786309, 0.897030549)$$

$$Y^{(3)} = (0.149121067, 0.273186438, 0.546647542, 0.844218268)$$

- We use the permutations:

$$\pi_1 = (1,2,3,4), \pi_2 = (2,1,3,4), \pi_3 = (3,4,2,1)$$

- This results in the sample

$$X_1 = (0.095500046, 0.349131748, 0.546647542)$$

$$X_2 = (0.493293558, 0.025170141, 0.844218268)$$

$$X_3 = (0.701216163, 0.705786309, 0.273186438)$$

$$X_4 = (0.866725669, 0.897030549, 0.149121067)$$

Matlab Codes Related to LHS

Matlab Codes

```
%%
clear all
n = 100; % Number of points
p = 2; % Number of variables
%real numbers between 0 and 1 that are drawn from a uniform distribution
x = rand(n,1);
y = rand(n,1);
subplot(2,3,1), plot(x,y,'o')
title('Simple Random Numbers(100 Points)')
% Latin hypercube sampling
X1 = lhsdesign(n,p);
subplot(2,3,2), plot(X1(:,1),X1(:,2),'o')
title('LHS(100 Points)')
% LHS with reducing correlation
X2 = lhsdesign(n,p,'criterion','correlation');
subplot(2,3,3), plot(X2(:,1),X2(:,2),'o')
title('LHS with reducing correlation(100 Points)')
% LHS with 5000 iterations to improve the design
X3 = lhsdesign(n,p,'iterations',5000);
subplot(2,3,4), plot(X3(:,1),X3(:,2),'o')
title('LHS with 5000 Iterations(100 Points)')
% LHS with reducing correlation and 5000 iterations
X4 = lhsdesign(n,p,'criterion','correlation','iterations',5000);
subplot(2,3,5), plot(X4(:,1),X4(:,2),'o')
title('LHS with Reducing Correlation and 5000 Iterations(100 Points)')
%% Correlation coefficients
cor1 = corrcoef(X1); cor2 = corrcoef(X2);
cor3 = corrcoef(X3); cor4 = corrcoef(X4);
cor1;cor2;cor3;cor4;
```

Matlab Codes Related to LHS

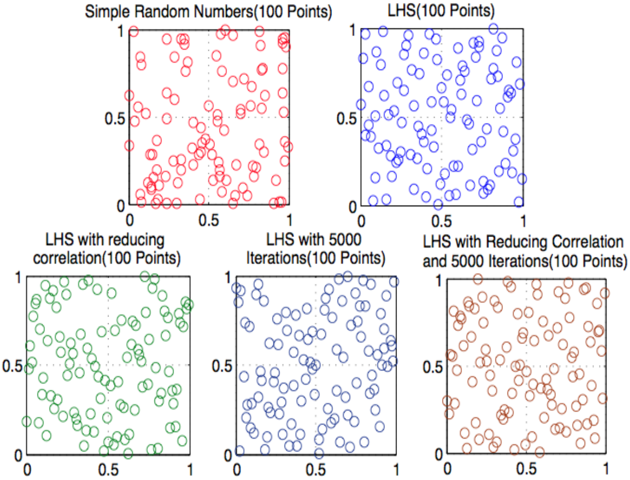


Figure: Latin Hypercube Sampling

Matlab Codes Related to LHS

Table: Correlation in Related Methods

Correlation	**Value**
LHS	-0.1255
LHS with reducing correlation	-0.0057
LHS with 5000 iterations	0.1618
LHS with reducing correlation and 5000 iterations	-0.0042

Matlab Codes Related to LHS

Matlab Codes

```
n=10000;     mu=0;     st=1;
for ii=1:100
         a=lhsnorm([mu mu],[st 0;0 st],n);
         x=randn(n,1);
    y=randn(n,1);
         ma(ii)=mean(a(:,1));
    mb(ii)=mean(a(:,2));
    mx(ii)=mean(x);     my(ii)=mean(y);
    t(ii)=ii;
end
mean(ma); mean(mb); mean(mx); mean(my)
```

Matlab Codes Related to LHS

Table: Means and STDs

	Mean	**STD**
LHS X	1.5822e-06	3.0788e-05
LHS Y	7.0585e-06	4.1981e-05
X	0.0111	0.9961
Y	0.0048	0.9960

Some Notes on LHS

- It is flexible in terms of data density and location, and it provides non-collapsing and space-filling properties.
- It has optimization of point location improving space filling and reducing correlation among points.
- In higher dimensions, it faces with problem unless K is small. Since size of at least K^d sample size is required to ensure each stratum is sampled.

Conditional Sampling

- Another variance reduction method similar to stratified sampling is the conditional sampling approach.
- Aim is to estimate $\mu = \mathbb{E}(X)$ and the variance reduction is achieved by another variable Y.
- In stratified sampling, the distribution of Y random variable is known; for instance we know the proportion of jobs in Ankara. Then we calculate the conditional expectation by Crude Monte Carlo.
- In conditional sampling, we know the conditional expectation, for instance through an analytical formula, but we do not know the distribution of Y.
- In other words, we do the reverse of Stratified Sampling.

Conditional Sampling

- Conditional expectation is calculated by the analytical formula.
- The distribution of Y is estimated by the crude Monte Carlo method through the following property:

$$\mu = \mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X|Y))$$

- Therefore, we firstly sample N times to obtain Y; $Y_1, Y_2, \ldots, Y_N$.
- Then we calculate the conditional expectation through an analytical formula.
- The last step is to average the conditional expectations.

$$X_{cond,N} = \frac{1}{N}\sum_{i=1}^{N} \mathbb{E}(X|Y = y_i)$$

Where does the Variance Reduction Come?

- $\sigma^2 = var(X) = \mathbb{E}(var(X|Y)) + var(\mathbb{E}(X|Y)) \geq var(\mathbb{E}(X|Y))$

Proof.

$$var(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 \quad (1)$$

$$var(X|Y) = \mathbb{E}(X^2|Y) - (\mathbb{E}(X|Y))^2 \quad (2)$$

$$\mathbb{E}(var(X|Y)) = \mathbb{E}(\mathbb{E}(X^2|Y)) - \mathbb{E}((\mathbb{E}(X|Y))^2) = \mathbb{E}(X^2) - \mathbb{E}((\mathbb{E}(X|Y))^2) \quad (3)$$

$$var(\mathbb{E}(X|Y)) = \mathbb{E}(\mathbb{E}(X|Y)^2) - (\mathbb{E}(\mathbb{E}(X|Y)))^2 = \mathbb{E}(\mathbb{E}(X|Y)^2) - \mathbb{E}(X)^2 \quad (4)$$

Sum the right hand side of Eq.3 and Eq.4, we reach the conclusion. □

Conditional Sampling

- The variance reduction through conditioning never exceeds the variance of the crude Monte Carlo estimator.
- If $X|Y$ is not almost surely constant, there is a positive variance reduction using the conditional Monte Carlo estimator.
- If the variance inside the groups is large and the conditional group means does not differ too much, we will have a big variance reduction.
- The confidence interval for conditional Monte Carlo is as follows:

$$\left[X_{cond,N} - 1.96\frac{\sigma_{cond}}{\sqrt{N}}, X_{cond,N} + 1.96\frac{\sigma_{cond}}{\sqrt{N}}\right]$$

Variance of Median for Normally Distributed Random Variables

- Suppose that we are trying to find out the variance of median for the standard normally distributed random variables $x_1, x_2, ..., x_n$.
- To apply Crude Monte Carlo, generate N samples of size n from $\mathcal{N}(0,1)$ and calculate the following:

$$\theta = \frac{1}{N}\sum_{j=1}^{N} M_j^2$$

- Alternative method is to use independence decomposition swindle. If X^{mean} is the sample mean,

$$var(M) = var\left(X^{mean} + (M - X^{mean})\right)$$

$$var(M) = var\left(X^{mean}\right) + var\left(M - X^{mean}\right) + 2cov\left(X^{mean}, (M - X^{mean})\right)$$

.

Variance of Median for Normally Distributed Random Variables

- However, X^{mean} is independent of $M - X^{mean}$, $cov\left(X^{mean}, (M - X^{mean})\right) = 0$.
- Then variance of median becomes as follows:

$$var(M) = var\left(X^{mean} + (M - X^{mean})\right)$$

$$var(M) = \frac{1}{n} + \frac{1}{N}\sum_{j=1}^{N}\left(M_j - X_j^{mean}\right)^2$$

- Therefore just use to calculate the variance of the difference between mean and median in conditional sampling Monte Carlo.

Variance of Median for Normally Distributed Random Variables

- Summarizing the results for the variance for median calculation.
- For samplesizes: 100 (1st) and 1000 (2nd), median's variance is as follows:

Number of Replicationas	Crude Monte Carlo	Conditional MC
100	0.1296	0.1187
1000	0.1245	0.1233
5000	0.1282	0.125
10000	0.1234	0.1243

Number of Replicationas	Crude Monte Carlo	Conditional MC
100	0.0421	0.0385
1000	0.0409	0.0398
5000	0.0397	0.0396
10000	0.0396	0.0396

- conditional MC's variance is generally lower than crude MC's

Application: Barrier Options

- A barrier option is a type of path-dependent option where the payoff is determined by whether or not the price of the stock crosses a certain level S_b during its life.
- There are two general types of barrier options, 'in' and 'out' options.
- In knock-out options, the contract is canceled if the barrier is crossed throughout the whole life.
- Knock-in options on the other hand are activated only if the barrier is crossed.
- For example, a down-and-out put option is a put option that becomes worthless if the asset price falls below the barrier S_b.
- It is reasonable to expect that a down-and-out put option is cheaper than a vanilla one, since it may expire worthless if the barrier is hit while the vanilla option would have paid off.

Barrier Options

- For instance, price of a down and out put option is as follows:

$$P = Ke^{rT}\left\{N(d_4) - N(d_2) - a\left(N(d_7) - N(d_5)\right)\right\} - S_0\left\{N(d_3) - N(d_1) - b\left(N(d_8) - N(d_6)\right)\right\}$$

where $a = \left\{\frac{S_a}{S_0}\right\}^{-1+\frac{2r}{\sigma^2}}$ and $b = \left\{\frac{S_b}{S_0}\right\}^{-1+\frac{2r}{\sigma^2}}$.

- d_i is function of S_0 (spot price of underlying), K (strike), r (interest rate), T (time to maturity), S_b (barrier level) and σ (volatility).
- $N(d_i)$ is the cumulative normal density of d_i.
- Price of a down and in put option is equal to the difference between the put option minus the down and out put option.

$$P = P_{down\&out} + P_{down\&in}$$

Application of Crude Monte Carlo for Barrier Options

- But barrier options are path-dependent options–the payoff is determined by whether or not the price of the asset hits a certain barrier during the life of the option.
- Due to this path-dependency, simulation of the entire price evolution is necessary.
- For instance, for down-and-out put option is in the money if the barrier is never reached. If not, the option is worthless.

Algorithm for Down-and-In Barrier Option-Crude MC

- The first step in pricing barrier options using Monte Carlo methods is to simulate the sample paths of the underlying stock price.
- To simulate a sample path, we have to choose a stochastic differential equation describing the dynamics of the price. We consider the price of the underlying asset is described by a Geometric Brownian Motion.
- $dS_t = rS_t dt + \sigma S_t dW_t$ and $S_{t+\Delta t} = S_t e^{(r-0.5\sigma^2)\Delta t + \sigma \epsilon_t \sqrt{\Delta t}}$
- If the path reaches the barrier, then the value of the option is zero.
- If the path does not reach the barrier, calculate the payoff of the put option at time T.
- Calculate the value of the option at time T for all paths and average them and discounted expected value is the price for the down and in put option.

Application of Conditional Sampling Technique for Barrier Options

- The first step in pricing barrier options using conditional expectation in Monte Carlo methods is to simulate the sample paths of the underlying stock price based on the stochastic process of the underlying.
- Here we calculate the price of down-and-in put option.
- If the path does not reach the barrier, then down-and-in put option is out of the money and its value is zero.
- If the path reaches the barrier, calculate the payoff of the put option using the analytical formula, since in this case the option becomes a European put option.
- Then average the final values for the option for each path, calculated by the analytical formula, and discount them.
- The last step is to deduct the down-and-in put option from European put option. Therefore, we reach the price of down-and-out put option.

Comparison of Results

- Under the example of $S_0 = 50, K = 50, r = 0.1, T = 1, S_b = 40$ and $\sigma = 0.2$; the actual price is 0.6264.
- The comparison of Crude Monte Carlo and conditional sampling is as follows:
  When the number of steps is 100 in each replication;

Table: Results

Crude Monte Carlo					Conditional Sampling			
Trials	Price	CI		Stand. Err.	Price	CI		Stand. Err.
100	0.7988	0.419	1.1787	0.19143	0.891	0.3569	1.4251	0.26917
1000	0.6997	0.5924	0.8071	0.054717	0.6471	0.466	0.8282	0.092291
5000	0.6984	0.6514	0.7454	0.023986	0.6892	0.609	0.7694	0.04092
10000	0.7149	0.6814	0.7483	0.017059	0.7136	0.6574	0.7697	0.028668
100000	0.7	0.6895	0.7106	0.005384	0.6948	0.6769	0.7126	0.009124

- For instance, for down-and-out put option is in the money if the barrier is never reached. If not, the option is worthless.

Comparison of Results

If the number of steps is 365 in each replication;

Table: Results

Crude Monte Carlo					**Conditional Sampling**			
Trials	Price	CI		Stand. Err.	Price	CI		Stand. Err.
100	0.726	1.6356	0.4014	1.0505	0.6831	2.8617	0.1153	1.2509
1000	0.7716	1.8092	0.6593	0.8838	0.6206	2.9359	0.4385	0.8028
5000	0.6787	1.6694	0.6324	0.7249	0.685	2.8433	0.6062	0.7639
10000	0.6565	1.6161	0.6248	0.6882	0.6407	2.8962	0.5839	0.6974
100000	0.6682	1.6507	0.658	0.6785	0.6687	2.8641	0.651	0.6865

We observe that the conditional expectation Monte Carlo estimator gives more accurate and precise estimates than the crude Monte Carlo estimator. However, the conditional expectation does not yield errors lower than that of the crude Monte Carlo.

Comparison of Results 2

- Under the example of $S_0 = 50, K = 50, r = 0.1, T = 1, S_b = 30$ and $\sigma = 0.2$; the actual price is 1.8136.
- The comparison of Crude Monte Carlo and conditional sampling is as follows:
  When the number of steps is 100 in each replication;

Table: Results

Crude Monte Carlo					**Conditional Sampling**			
Trials	Price	CI		Stand. Err.	Price	CI		Stand. Err.
100	2.3873	1.6181	3.1566	0.1226	1.7723	1.1264	2.4182	0.3255
1000	1.7565	1.5558	1.9572	0.1023	1.8029	1.7306	1.8752	0.0368
5000	1.8025	1.7100	1.8950	0.0472	1.8295	1.8039	1.8551	0.0131
10000	1.8154	1.7493	1.8815	0.0337	1.8410	1.8253	1.8566	0.0080
100000	1.8285	1.8074	1.8495	0.0107	1.8228	1.8167	1.8289	0.0031

Comparison of Results 2

- If the number of steps is 365 in each replication;

Table: Results

Crude Monte Carlo					Conditional Sampling			
Trials	Price	CI		Stand. Err.	Price	CI		Stand. Err.
100	1.3198	0.7808	1.8588	0.2716	1.5118	1.0121	2.0196	0.2539
1000	1.8151	1.5980	2.0322	0.1106	1.8053	1.7354	1.8753	0.0356
5000	1.8908	1.7945	1.9871	0.0491	1.8128	1.7833	1.8423	0.0150
10000	1.8464	1.7794	1.9135	0.0342	1.8078	1.7862	1.8294	0.0110
100000	1.8081	1.7872	1.8289	0.0106	1.8183	1.8120	1.8246	0.0032

We observe that when the barrier level is decreased, in other words, when the distance between barrier level and current underlying stock price increases, conditional MC gives us smaller standard errors compared to Crude MC.

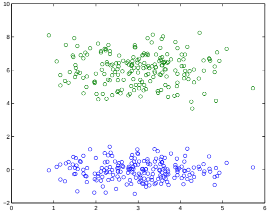
Summary and Comparison

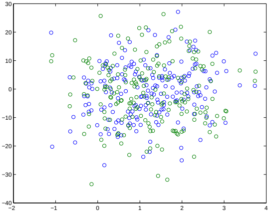
Conditional Sampling	Stratified Sampling
Y estimated by MC	Y is known
$\mathbb{E}(X\|Y)$ achieved analytically (known)	$X\|Y$ estimated by MC
$\mu = \mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X\|Y))$	$\mu = \mathbb{E}(X) = \sum_{i=1}^{d} \mathbb{E}(X\|Y = y_i)\ P(Y = y_i)$

Variance Decomposition

$Var(X) = \mathbb{E}(Var(X|Y)) + Var(\mathbb{E}(X|Y)) \geq \mathbb{E}(Var(X|Y)) = N\sigma^2_{prop-strat}$
whereas value of $Var(\mathbb{E}(X|Y))$ depicts reduction of variance compared to crude MC for stratified estimator.

$Var(X) = \mathbb{E}(Var(X|Y)) + Var(\mathbb{E}(X|Y)) \geq Var(\mathbb{E}(X|Y)) = \sigma^2_{cond}$
whereas value of $\mathbb{E}(Var(X|Y))$ depicts reduction of variance compared to crude MC for conditional estimator.





stratified sampling should be applied | conditional sampling should be applied

Confidence Intervals - Comparison

- $\left[\overline{X}_{crude,N} - \frac{z_{1-\alpha/2}}{\sqrt{N}}\sigma \ , \ \overline{X}_{crude,N} + \frac{z_{1-\alpha/2}}{\sqrt{N}}\sigma\right]$
- $\left[\overline{X}_{strat,N} - \frac{z_{1-\alpha/2}}{\sqrt{N}}\sqrt{\sum_{i=1}^{d}\hat{\sigma}_i^2 p_i} \ , \ \overline{X}_{strat,N} + \frac{z_{1-\alpha/2}}{\sqrt{N}}\sqrt{\sum_{i=1}^{d}\hat{\sigma}_i^2 p_i}\right]$
- $\left[\overline{X}_{cond,N} - \frac{z_{1-\alpha/2}}{\sqrt{N}}\sigma_{cond} \ , \ \overline{X}_{cond,N} + \frac{z_{1-\alpha/2}}{\sqrt{N}}\sigma_{cond}\right]$