Introduction to Simulation Methods,

Michelle Sovinsky Goeree

University of Zuerich

October 17, 2011

Outline

- Generating Random Variables
- Simulation of Integrals
- Variance Reduction
- Simulation in Estimation

- Simulation methods can be used for demand estimation, option pricing, risk , econometrics, etc.
- Naive Monte Carlo may be too slow in some practical situations.
- Many special techniques for variance reduction: antithetic variables, importance sampling, etc.

- Simulation methods can be used for demand estimation, option pricing, risk , econometrics, etc.
- Naive Monte Carlo may be too slow in some practical situations.
- Many special techniques for variance reduction: antithetic variables, importance sampling, etc.

- Simulation methods can be used for demand estimation, option pricing, risk , econometrics, etc.
- Naive Monte Carlo may be too slow in some practical situations.
- Many special techniques for variance reduction: antithetic variables, importance sampling, etc.

The Basics

- Consider the basic problem of computing an expectation

$$\theta = E[f(X)], \qquad X \sim pdf(X)$$

- Monte Carlo simulation approach specifies generating N independent draws from the distribution $pdf(X), X_1, X_2, ..., X_N$, and approximating

$$E[f(X)] \approx \widehat{\theta}_N \equiv \frac{1}{N} \sum_{i=1}^{N} f(X_i)$$

- By Law of Large Numbers, the approximation $\widehat{\theta}_N$ converges to the true value as N increases to infinity.

The Basics

- Consider the basic problem of computing an expectation

$$\theta = E[f(X)], \qquad X \sim pdf(X)$$

- Monte Carlo simulation approach specifies generating N independent draws from the distribution $pdf(X), X_1, X_2, ..., X_N$, and approximating

$$E[f(X)] \approx \widehat{\theta}_N \equiv \frac{1}{N}\sum_{i=1}^{N} f(X_i)$$

- By Law of Large Numbers, the approximation $\widehat{\theta}_N$ converges to the true value as N increases to infinity.

The Basics

- Consider the basic problem of computing an expectation

$$\theta = E[f(X)], \qquad X \sim pdf(X)$$

- Monte Carlo simulation approach specifies generating N independent draws from the distribution $pdf(X), X_1, X_2, ..., X_N$, and approximating

$$E[f(X)] \approx \widehat{\theta}_N \equiv \frac{1}{N}\sum_{i=1}^{N} f(X_i)$$

- By Law of Large Numbers, the approximation $\widehat{\theta}_N$ converges to the true value as N increases to infinity.

The Basics

- Monte Carlo estimate $\widehat{\theta}_N$ is unbiased:

$$E[\widehat{\theta}_N] = \theta$$

- By Central Limit Theorem
$\sqrt{N}\frac{\widehat{\theta}_N - \theta}{\sigma} \Rightarrow N(0,1), \sigma^2 = Var[f(X)]$

The Basics

- Monte Carlo estimate $\widehat{\theta}_N$ is unbiased:

$$E[\widehat{\theta}_N] = \theta$$

- By Central Limit Theorem
$\sqrt{N}\frac{\widehat{\theta}_N - \theta}{\sigma} \Rightarrow N(0,1), \sigma^2 = Var[f(X)]$

- Pseudo random number generators produce deterministic sequences of numbers that appear stochastic, and match closely the desired probability distribution.
- For some standard distributions, e.g., uniform and Normal, MATLAB® provides built-in random number generators.
- Sometimes it is necessary to simulate from other distributions, not covered by the standard software.
- Then apply one of the basic methods for generating random variables from a specified distribution.

- Pseudo random number generators produce deterministic sequences of numbers that appear stochastic, and match closely the desired probability distribution.
- For some standard distributions, e.g., uniform and Normal, MATLAB® provides built-in random number generators.
- Sometimes it is necessary to simulate from other distributions, not covered by the standard software.
- Then apply one of the basic methods for generating random variables from a specified distribution.

- Pseudo random number generators produce deterministic sequences of numbers that appear stochastic, and match closely the desired probability distribution.
- For some standard distributions, e.g., uniform and Normal, MATLAB® provides built-in random number generators.
- Sometimes it is necessary to simulate from other distributions, not covered by the standard software.
- Then apply one of the basic methods for generating random variables from a specified distribution.

The Inverse Transform Method

- Consider a random variable X with a continuous, strictly increasing CDF function $F(x)$.
- We can simulate X according to

$$X = F^{-1}(U), \qquad U \sim Unif[0,1]$$

- This works because

$$\Pr(X \leq x) = \Pr(F^{-1}(U) \leq x) = \Pr(U \leq F(x)) = F(x)$$

- If $F(x)$ has jumps or flat sections, generalize the above rule to

$$X = \min(x : F(x) \geq U)$$

The Inverse Transform Method

- Consider a random variable X with a continuous, strictly increasing CDF function $F(x)$.
- We can simulate X according to

$$X = F^{-1}(U), \qquad U \sim Unif[0,1]$$

- This works because

$$\Pr(X \leq x) = \Pr(F^{-1}(U) \leq x) = \Pr(U \leq F(x)) = F(x)$$

- If $F(x)$ has jumps or flat sections, generalize the above rule to

$$X = \min(x : F(x) \geq U)$$

The Inverse Transform Method

- Consider a random variable X with a continuous, strictly increasing CDF function $F(x)$.
- We can simulate X according to

$$X = F^{-1}(U), \qquad U \sim Unif[0,1]$$

- This works because

$$\Pr(X \leq x) = \Pr(F^{-1}(U) \leq x) = \Pr(U \leq F(x)) = F(x)$$

- If $F(x)$ has jumps or flat sections, generalize the above rule to

$$X = \min(x : F(x) \geq U)$$

The Inverse Transform Method

- Consider a random variable X with a continuous, strictly increasing CDF function $F(x)$.
- We can simulate X according to

$$X = F^{-1}(U), \qquad U \sim Unif[0,1]$$

- This works because

$$\Pr(X \leq x) = \Pr(F^{-1}(U) \leq x) = \Pr(U \leq F(x)) = F(x)$$

- If $F(x)$ has jumps or flat sections, generalize the above rule to

$$X = \min(x : F(x) \geq U)$$

Example: The Exponential Distribution

- Consider an exponentially-distributed random variable, characterized by a CDF

$$F(x) = U = 1 - e^{-x\theta}$$

- Compute $F^{-1}(U)$

$$e^{-x\theta} = 1 - U \Rightarrow \ln(1 - U) = -x\theta$$

- So

$$X = \frac{-\ln(1-U)}{\theta} \sim \frac{-\ln(U)}{\theta}$$

Example: The Exponential Distribution

- Consider an exponentially-distributed random variable, characterized by a CDF

$$F(x) = U = 1 - e^{-x\theta}$$

- Compute $F^{-1}(U)$

$$e^{-x\theta} = 1 - U \Rightarrow \ln(1 - U) = -x\theta$$

- So

$$X = \frac{-\ln(1 - U)}{\theta} \sim \frac{-\ln(U)}{\theta}$$

Example: Normal Distribution

$$X \sim N(\mu, \sigma^2)$$

$$F(x) = U = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

$$\sigma\Phi^{-1}(U) + \mu = X \sim N(\mu, \sigma^2)$$

Example: Normal Distribution

$$X \sim N(\mu, \sigma^2)$$

$$F(x) = U = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

$$\sigma\Phi^{-1}(U) + \mu = X \sim N(\mu, \sigma^2)$$

Example: Discrete Distribution

- Consider a discrete random variable X with values

$$c_1 < c_2 < \cdots < c_n$$

$$\Pr(X = c_i) = p$$

- Define cummulative proababilities

$$F(c_i) = q_i = \sum_{i=1}^{N} p_i$$

- Can simulate X as follows:
  1. Generate $U \sim Unif[0, 1]$
  2. Find $K \in \{1, ..., n\}$ such that $q_{K-1} \leq U \leq q_K$
  3. Set $X = c_K$

Example: Discrete Distribution

- Consider a discrete random variable X with values

$$c_1 < c_2 < \cdots < c_n$$

$$\Pr(X = c_i) = p$$

- Define cummulative proabablities

$$F(c_i) = q_i = \sum_{i=1}^{N} p_i$$

- Can simulate X as follows:
  1. Generate $U \sim Unif[0, 1]$
  2. Find $K \in \{1, ..., n\}$ such that $q_{K-1} \leq U \leq q_K$
  3. Set $X = c_K$

Example: Discrete Distribution

- Consider a discrete random variable X with values

$$c_1 < c_2 < \cdots < c_n$$

$$\Pr(X = c_i) = p$$

- Define cummulative proababilities

$$F(c_i) = q_i = \sum_{i=1}^{N} p_i$$

- Can simulate X as follows:
  1. Generate $U \sim Unif[0, 1]$
  2. Find $K \in \{1, ..., n\}$ such that $q_{K-1} \leq U \leq q_K$
  3. Set $X = c_K$

Examples

- Bernouli where $X = 1$ with probability p and 0 with probability $1 - p$
$$X = 1(U < p)$$
- Binomial
$$X \sim Bin(N, p)$$

1. $$X = F^{-1}(U)$$
2. $$X = \sum_{i=1}^{N} 1(U_i < p)$$

Examples

- Bernouli where $X = 1$ with probability p and 0 with probability $1 - p$

$$X = 1(U < p)$$

- Binomial

$$X \sim Bin(N, p)$$

1.
$$X = F^{-1}(U)$$

2.
$$X = \sum_{i=1}^{N} 1(U_i < p)$$

Examples

- Bernouli where $X = 1$ with probability p and 0 with probability $1 - p$
$$X = 1(U < p)$$
- Binomial
$$X \sim Bin(N, p)$$

1.
$$X = F^{-1}(U)$$

2.
$$X = \sum_{i=1}^{N} 1(U_i < p)$$

Examples

- Bernouli where $X = 1$ with probability p and 0 with probability $1 - p$
$$X = 1(U < p)$$
- Binomial

$$X \sim Bin(N, p)$$

1. $$X = F^{-1}(U)$$

2. $$X = \sum_{i=1}^{N} 1(U_i < p)$$

Examples

- Chi-Square

$$X \sim \chi_n^2$$

1. $$X = F^{-1}(U)$$

2. $$X = \sum_{i=1}^{N} [\Phi^{-1}(U)]^2$$

Examples

- Chi-Square

$$X \sim \chi^2_n$$

1.

$$X = F^{-1}(U)$$

2.

$$X = \sum_{i=1}^{N} [\Phi^{-1}(U)]^2$$

The Acceptance-Rejection Method

- Generate samples with probability density $f(x)$
- The acceptance-rejection method can be used for multivariate problems as well
- Suppose we know how to generate samples from the distribution with pdf $g(x)$ such that $f(x) \leq cg(x)$
- Follow the algorithm

1. Generate X from the distribution $g(x)$
2. Generate U from $Unif[0, 1]$
3. If $U \leq f(x)/[cg(X)]$, return X, otherwise go to step 1

- Probability of acceptance on each attempt is $1/c$. Want c close to 1

The Acceptance-Rejection Method

- Generate samples with probability density $f(x)$
- The acceptance-rejection method can be used for multivariate problems as well
- Suppose we know how to generate samples from the distribution with pdf $g(x)$ such that $f(x) \leq cg(x)$
- Follow the algorithm

1. Generate X from the distribution $g(x)$
2. Generate U from $Unif[0,1]$
3. If $U \leq f(x)/[cg(X)]$, return X, otherwise go to step 1

- Probability of acceptance on each attempt is $1/c$. Want c close to 1

The Acceptance-Rejection Method

- Generate samples with probability density $f(x)$
- The acceptance-rejection method can be used for multivariate problems as well
- Suppose we know how to generate samples from the distribution with pdf $g(x)$ such that $f(x) \leq cg(x)$
- Follow the algorithm

1. Generate X from the distribution $g(x)$
2. Generate U from $Unif[0,1]$
3. If $U \leq f(x)/[cg(X)]$, return X, otherwise go to step 1

- Probability of acceptance on each attempt is $1/c$. Want c close to 1

Example: Truncated Random Variables

$$X \sim F : a < X < b$$

1. Acceptance-Rejection

$$X = F^{-1}(U)$$

   Keep if $a < X < b$ otherwise tray again

2.

$$U = \frac{F(X) - F(a)}{F(b) - F(a)}$$

$$F^{-1}\{[F(b) - F(a)]U + F(a)\} = X$$

Example: Truncated Random Variables

$$X \sim F : a < X < b$$

1. Acceptance-Rejection

$$X = F^{-1}(U)$$

   Keep if $a < X < b$ otherwise tray again

2.

$$U = \frac{F(X) - F(a)}{F(b) - F(a)}$$

$$F^{-1}\{[F(b) - F(a)]U + F(a)\} = X$$

Example: Truncated Random Variables

$$X \sim F : a < X < b$$

1. Acceptance-Rejection

$$X = F^{-1}(U)$$

   Keep if $a < X < b$ otherwise tray again

2. $$U = \frac{F(X) - F(a)}{F(b) - F(a)}$$

$$F^{-1}\{[F(b) - F(a)]U + F(a)\} = X$$

Example: Truncated Random Variables

- Consider the problem of computing

$$Eh(U) = \int h(u)dF(u)$$

- Moments in 'Method of Moments' Estimation
- Probabilities in discrete choice models

Example: Truncated Random Variables

- Consider the problem of computing

$$Eh(U) = \int h(u)dF(u)$$

- Moments in 'Method of Moments' Estimation
- Probabilities in discrete choice models

Prototypical Example

$$
\begin{aligned}
y_{ij}^{*} &= X_i \beta_j + u_{ij} \\
u_i &\sim iidF
\end{aligned}
$$

Prototypical Example

$$
\begin{aligned}
y_{ij}^{*} &= X_i \beta_j + u_{ij} \\
u_i &\sim iidF
\end{aligned}
$$

- $u_{ij} \sim iidEV \Rightarrow$ Multinomial Logit w/ benefits and disadvantages

Prototypical Example

$$\begin{aligned} y_{ij}^* &= X_i \beta_j + u_{ij} \\ u_i &\sim iidF \end{aligned}$$

- $u_{ij} \sim iidEV \Rightarrow$ Multinomial Logit w/ benefits and disadvantages
- $u_i \sim iidN(0, \Omega) \Rightarrow$ Multinomial Probit w/ benefits and disadvantages

Prototypical Example

$$\begin{aligned} y_{ij}^* &= X_i \beta_j + u_{ij} \\ u_i &\sim iidF \end{aligned}$$

- $u_{ij} \sim iidEV \Rightarrow$ Multinomial Logit w/ benefits and disadvantages
- $u_i \sim iidN(0, \Omega) \Rightarrow$ Multinomial Probit w/ benefits and disadvantages

$$\begin{aligned} U &\sim N(\mu, \Omega) \\ & \Pr[U < v] \end{aligned}$$

Basic Method

$$
\begin{aligned}
Z &= \frac{1}{R}\sum_{r=1}^{R} h\left(u^{r}\right) \\
u^{r} &\sim F
\end{aligned}
$$

Basic Method

$$EZ = E\left[\frac{1}{R}\sum_{r=1}^{R} h\left(u^r\right)\right]$$

Basic Method

$$
\begin{aligned}
EZ &= E\left[\frac{1}{R}\sum_{r=1}^{R} h\left(u^{r}\right)\right] \\
&= \frac{1}{R}\sum_{r=1}^{R} Eh\left(u^{r}\right)
\end{aligned}
$$

Basic Method

$$
\begin{aligned}
EZ &= E\left[\frac{1}{R}\sum_{r=1}^{R} h\left(u^{r}\right)\right] \\
&= \frac{1}{R}\sum_{r=1}^{R} Eh\left(u^{r}\right) \\
&= \frac{1}{R}\sum_{r=1}^{R} Eh\left(U\right) = Eh\left(U\right)
\end{aligned}
$$

Basic Method

$$VarZ = Var\left[\frac{1}{R}\sum_{r=1}^{R} h\left(u^{r}\right)\right]$$

Basic Method

$$VarZ = Var\left[\frac{1}{R}\sum_{r=1}^{R} h\left(u^r\right)\right]$$

$$= \frac{1}{R^2}\sum_{r=1}^{R} Var\left[h\left(u^r\right)\right]$$

Basic Method

$$VarZ = Var\left[\frac{1}{R}\sum_{r=1}^{R} h\left(u^{r}\right)\right]$$

$$= \frac{1}{R^2}\sum_{r=1}^{R} Var\left[h\left(u^{r}\right)\right]$$

$$= \frac{1}{R^2}\sum_{r=1}^{R} Var\left[h\left(U\right)\right] = \frac{Var\left[h\left(U\right)\right]}{R}$$

Basic Method for Multivariate Normal Probability

$$X \sim N(\mu, \Omega)$$

$$\Pr[X < v]$$

Basic Method for Multivariate Normal Probability

$$
\begin{aligned}
X &\sim N(\mu, \Omega) \\
& \Pr[X < v]
\end{aligned}
$$

$$h(X) = 1(X < v)$$

Basic Method for Multivariate Normal Probability

$$
\begin{aligned}
X &\sim N(\mu, \Omega) \\
& \Pr[X < v]
\end{aligned}
$$

$$h(X) = 1(X < v)$$

$$Eh(X) = \int 1(x < v)\, d\Phi(x; \mu, \Omega)$$

Basic Method for Multivariate Normal Probability

$$\begin{aligned} X &\sim N(\mu, \Omega) \\ &\Pr[X < v] \end{aligned}$$

$$h(X) = 1(X < v)$$

$$Eh(X) = \int 1(x < v)\, d\Phi(x; \mu, \Omega)$$

Frequency Method (Lerman & Manski)

$$\begin{aligned} Z &= \frac{1}{R} \sum_{r=1}^{R} 1(x^r < v) \\ x^r &\sim iidN(\mu, \Omega) \end{aligned}$$

Problems w Frequency Method

1. Not continuous in v

Problems w Frequency Method

1. Not continuous in v
2. Not bounded away from zero and one [problems for MLE]

Problems w Frequency Method

1. Not continuous in v
2. Not bounded away from zero and one [problems for MLE]
3. Unnecessarily large variance

Improvements

Importance Sampling

$$Eh(U) = \int h(u) f(u) du$$

Improvements

Importance Sampling

$$Eh(U) = \int h(u) f(u) \, du$$

$$= \int \frac{h(u) f(u)}{g(u)} g(u) \, du$$

$$= E_g \left[\frac{h(U) f(U)}{g(U)} \right]$$

Improvements

Importance Sampling

$$Eh(U) = \int h(u) f(u) \, du$$

$$\begin{aligned} &= \int \frac{h(u) f(u)}{g(u)} g(u) \, du \\ &= E_g \left[\frac{h(U) f(U)}{g(U)} \right] \end{aligned}$$

Properties of a Good Importance Sampler

1. Support of G = support of F

Improvements

Importance Sampling

$$Eh(U) = \int h(u) f(u) \, du$$

$$= \int \frac{h(u) f(u)}{g(u)} g(u) \, du$$

$$= E_g \left[\frac{h(U) f(U)}{g(U)} \right]$$

Properties of a Good Importance Sampler

1. Support of G = support of F
2. Easy to simulate from G

Improvements

Importance Sampling

$$Eh(U) = \int h(u) f(u)\, du$$

$$= \int \frac{h(u) f(u)}{g(u)} g(u)\, du$$
$$= E_g\left[\frac{h(U) f(U)}{g(U)}\right]$$

Properties of a Good Importance Sampler

1. Support of G = support of F
2. Easy to simulate from G
3. $\frac{h(U)f(U)}{g(U)}$ does not vary as much as $h(U)$

Improvements

Importance Sampling

$$\begin{aligned} Eh(U) &= \int 1(u < v) f(u)\, du \\ &= \int_{u<v} f(u)\, du \end{aligned}$$

Improvements
Importance Sampling

$$
\begin{aligned}
Eh(U) &= \int 1(u < v) f(u)\, du \\
&= \int_{u<v} f(u)\, du
\end{aligned}
$$

Let G be independent truncated normals

Improvements

Importance Sampling

$$
\begin{aligned}
Eh(U) &= \int 1(u < v) f(u)\, du \\
&= \int_{u<v} f(u)\, du
\end{aligned}
$$

Let G be independent truncated normals

1. Support of G = support of F

Improvements

Importance Sampling

$$
\begin{aligned}
Eh(U) &= \int 1(u < v) f(u)\, du \\
&= \int_{u<v} f(u)\, du
\end{aligned}
$$

Let G be independent truncated normals

1. Support of G = support of F
2. Easy to simulate from G

Improvements

Importance Sampling

$$\begin{aligned} Eh(U) &= \int 1(u < v) f(u) \, du \\ &= \int_{u<v} f(u) \, du \end{aligned}$$

Let G be independent truncated normals

1. Support of G = support of F
2. Easy to simulate from G
3. $\frac{h(U)f(U)}{g(U)}$ is unbounded

Improvements

GHK

Algorithm

1. Initialize $P = 1$
2. Compute $\Pr(U_1 < v_1)$, and update $P = P * \Pr(U_1 < v_1)$

Improvements
GHK

Algorithm

1. Initialize $P = 1$
2. Compute $\Pr(U_1 < v_1)$, and update $P = P * \Pr(U_1 < v_1)$
3. Simulate $U_1^r \mid U_1^r < v_1$

Improvements
GHK

Algorithm

1. Initialize $P = 1$
2. Compute $\Pr(U_1 < v_1)$, and update $P = P * \Pr(U_1 < v_1)$
3. Simulate $U_1^r \mid U_1^r < v_1$
4. Initialize $i = 1$
5. Update $i \Leftarrow i + 1$

Improvements

GHK

Algorithm

1. Initialize $P = 1$
2. Compute $\Pr(U_1 < v_1)$, and update $P = P * \Pr(U_1 < v_1)$
3. Simulate $U_1^r \mid U_1^r < v_1$
4. Initialize $i = 1$
5. Update $i \Leftarrow i + 1$
6. Compute density of $U_i \mid U_1^r, U_2^r, .., U_{i-1}^r$

Improvements
GHK

Algorithm

1. Initialize $P = 1$
2. Compute $\Pr\left(U_1 < v_1\right)$, and update $P = P * \Pr\left(U_1 < v_1\right)$
3. Simulate $U_1^r \mid U_1^r < v_1$
4. Initialize $i = 1$
5. Update $i \Leftarrow i + 1$
6. Compute density of $U_i \mid U_1^r, U_2^r, .., U_{i-1}^r$
7. Compute $\Pr\left(U_i < v_i \mid U_1^r, U_2^r, .., U_{i-1}^r\right)$, and update $P = P * \Pr\left(U_i < v_i \mid U_1^r, U_2^r, .., U_{i-1}^r\right)$

Improvements
GHK

Algorithm

1. Initialize $P = 1$
2. Compute $\Pr(U_1 < v_1)$, and update $P = P * \Pr(U_1 < v_1)$
3. Simulate $U_1^r \mid U_1^r < v_1$
4. Initialize $i = 1$
5. Update $i \Leftarrow i + 1$
6. Compute density of $U_i \mid U_1^r, U_2^r, .., U_{i-1}^r$
7. Compute $\Pr\left(U_i < v_i \mid U_1^r, U_2^r, .., U_{i-1}^r\right)$, and update $P = P * \Pr\left(U_i < v_i \mid U_1^r, U_2^r, .., U_{i-1}^r\right)$
8. Return to (5) until done

Improvements
GHK

1. Support of G = support of F

Improvements
GHK

1. Support of G = support of F
2. Easy to simulate from G

Improvements
GHK

1. Support of G = support of F
2. Easy to simulate from G
3. Simulator is bounded away from zero and one, and variance is smaller than frequency simulator

Anithetic Acceleration

- Attempt to reduce variance due to simulation
- Introduce negative dependence between pairs of replications

Improvements

Antithetic Acceleration

Let $U \sim U(0,1)$, and consider as a simulator for $Eh(U)$,

$$S = \frac{1}{2R} \sum_{r=1}^{2R} h(u^r)$$

Improvements

Antithetic Acceleration

Let $U \sim U(0,1)$, and consider as a simulator for $Eh(U)$,

$$S = \frac{1}{2R} \sum_{r=1}^{2R} h(u^r)$$

Now compare it to

$$AS = \frac{1}{2R} \sum_{r=1}^{2R} [h(u^r) + h(1 - u^r)]$$

Improvements

Antithetic Acceleration

Let $U \sim U(0,1)$, and consider as a simulator for $Eh(U)$,

$$S = \frac{1}{2R} \sum_{r=1}^{2R} h(u^r)$$

Now compare it to

$$AS = \frac{1}{2R} \sum_{r=1}^{2R} [h(u^r) + h(1 - u^r)]$$

$$\begin{aligned} E(AS) &= \frac{1}{2R} \sum_{r=1}^{2R} [Eh(u^r) + Eh(1 - u^r)] \\ &= \frac{1}{2R} \sum_{r=1}^{2R} [Eh(U) + Eh(1 - U)] \end{aligned}$$

Improvements

Antithetic Acceleration

Let $U \sim U(0,1)$, and consider as a simulator for $Eh(U)$,

$$S = \frac{1}{2R} \sum_{r=1}^{2R} h(u^r)$$

Now compare it to

$$AS = \frac{1}{2R} \sum_{r=1}^{2R} \left[h(u^r) + h(1-u^r) \right]$$

$$\begin{aligned} E(AS) &= \frac{1}{2R} \sum_{r=1}^{2R} \left[Eh(u^r) + Eh(1-u^r) \right] \\ &= \frac{1}{2R} \sum_{r=1}^{2R} \left[Eh(U) + Eh(1-U) \right] \end{aligned}$$

$$= \frac{1}{2R} \sum_{r=1}^{2R} \left[Eh(U) + Eh(U) \right] = Eh(U)$$

Improvements

Antithetic Acceleration

$$\begin{aligned} Var\left(AS\right) &= \frac{1}{4R^2}\sum_{r=1}^{2R} Var\left[h\left(u^r\right)+h\left(1-u^r\right)\right] \\ &= \frac{1}{4R^2}\sum_{r=1}^{2R}\left[Var\left(h\left(u^r\right)\right)+Var\left(h\left(1-u^r\right)\right)\right. \\ &\quad \left. +2Cov\left(h\left(u^r\right),h\left(1-u^r\right)\right)\right] \end{aligned}$$

Improvements

Antithetic Acceleration

$$
\begin{aligned}
Var(AS) &= \frac{1}{4R^2}\sum_{r=1}^{2R} Var\left[h(u^r) + h(1-u^r)\right] \\
&= \frac{1}{4R^2}\sum_{r=1}^{2R} \left[Var(h(u^r)) + Var(h(1-u^r))\right. \\
&\quad \left. +2Cov(h(u^r), h(1-u^r))\right]
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{4R^2}\sum_{r=1}^{2R} \left[Var(h(U)) + Var(h(1-U))\right. \\
&\quad \left. +2Cov(h(U), h(1-U))\right] \\
&= \frac{1}{2R}\sum_{r=1}^{2R} \left[Var(h(U)) + Cov(h(U), h(1-U))\right]
\end{aligned}
$$

Improvements

Antithetic Acceleration

$$Var\left(AS\right) = \frac{1}{2R}\sum_{r=1}^{2R}\left[Var\left(h\left(U\right)\right) + Cov\left(h\left(U\right), h\left(1-U\right)\right)\right]$$

- If $Cov\left(h\left(U\right), h\left(1-U\right)\right) = 0$, then $Var\left(AS\right) = Var\left(S\right)$

Improvements

Antithetic Acceleration

$$Var\left(AS\right) = \frac{1}{2R}\sum_{r=1}^{2R}\left[Var\left(h\left(U\right)\right) + Cov\left(h\left(U\right), h\left(1-U\right)\right)\right]$$

- If $Cov\left(h\left(U\right), h\left(1-U\right)\right) = 0$, then $Var\left(AS\right) = Var\left(S\right)$
- If $Cov\left(h\left(U\right), h\left(1-U\right)\right) < 0$, then $Var\left(AS\right) < Var\left(S\right)$

Improvements

Antithetic Acceleration

$$Var\left(AS\right)=\frac{1}{2R}\sum_{r=1}^{2R}\left[Var\left(h\left(U\right)\right)+Cov\left(h\left(U\right),h\left(1-U\right)\right)\right]$$

- If $Cov\left(h\left(U\right),h\left(1-U\right)\right)=0$, then $Var\left(AS\right)=Var\left(S\right)$
- If $Cov\left(h\left(U\right),h\left(1-U\right)\right)<0$, then $Var\left(AS\right)<Var\left(S\right)$ [A sufficient condition is $h\left(\cdot\right)$ is monotone]

Improvements

Antithetic Acceleration

$$Var\left(AS\right)=\frac{1}{2R}\sum_{r=1}^{2R}\left[Var\left(h\left(U\right)\right)+Cov\left(h\left(U\right),h\left(1-U\right)\right)\right]$$

- If $Cov\left(h\left(U\right),h\left(1-U\right)\right)=0$, then $Var\left(AS\right)=Var\left(S\right)$
- If $Cov\left(h\left(U\right),h\left(1-U\right)\right)<0$, then $Var\left(AS\right)<Var\left(S\right)$ [A sufficient condition is $h\left(\cdot\right)$ is monotone]
- If $Cov\left(h\left(U\right),h\left(1-U\right)\right)>0$, then $Var\left(AS\right)>Var\left(S\right)$

MSM

Many estimation problems can be framed as

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^{n} Z_i' e_i\left(\widehat{\theta}_{MOM}\right) &= 0 \\ e_i(\theta) &= y_i - E\left[y_i \mid X_i, \theta\right] \end{aligned}$$

MSM

Many estimation problems can be framed as

$$\frac{1}{n}\sum_{i=1}^{n} Z_i' e_i\left(\widehat{\theta}_{MOM}\right) = 0$$

$$e_i(\theta) = y_i - E\left[y_i \mid X_i, \theta\right]$$

Sometimes it is difficult to evaluate $E\left[y_i \mid X_i, \theta\right]$ while it is not difficult to simulate $E\left[y_i \mid X_i, \theta\right]$

MSM

Many estimation problems can be framed as

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^{n} Z_i' e_i\left(\widehat{\theta}_{MOM}\right) &= 0 \\ e_i(\theta) &= y_i - E\left[y_i \mid X_i, \theta\right] \end{aligned}$$

Sometimes it is difficult to evaluate $E\left[y_i \mid X_i, \theta\right]$ while it is not difficult to simulate $E\left[y_i \mid X_i, \theta\right]$

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^{n} Z_i' e_i\left(\widehat{\theta}_{MSM}\right) &= 0 \\ e_i(\theta) &= y_i - \widetilde{E}\left[y_i \mid X_i, \theta\right] \end{aligned}$$

where $\widetilde{E}\left[y_i \mid X_i, \theta\right]$ is an unbiased simulator of $E\left[y_i \mid X_i, \theta\right]$

MSM

$$
\begin{aligned}
\widetilde{e}_i(\theta) &= e_i(\theta) + \xi_i \\
E\xi_i &= 0
\end{aligned}
$$

MSM

$$
\begin{aligned}
\widetilde{e}_i(\theta) &= e_i(\theta) + \xi_i \\
E\xi_i &= 0
\end{aligned}
$$

$$
plim\left[\frac{1}{n}\sum_{i=1}^{n} Z_i' \widetilde{e}_i(\theta)\right] = plim\left[\frac{1}{n}\sum_{i=1}^{n} Z_i' \left(e_i(\theta) + \xi_i\right)\right]
$$

MSM

$$\begin{aligned} \widetilde{e}_i(\theta) &= e_i(\theta) + \xi_i \\ E\xi_i &= 0 \end{aligned}$$

$$plim\left[\frac{1}{n}\sum_{i=1}^{n} Z_i' \widetilde{e}_i(\theta)\right] = plim\left[\frac{1}{n}\sum_{i=1}^{n} Z_i' \left(e_i(\theta) + \xi_i\right)\right]$$

$$= plim\left[\frac{1}{n}\sum_{i=1}^{n} Z_i' e_i(\theta) + \frac{1}{n}\sum_{i=1}^{n} Z_i' \xi_i\right]$$

MSM

$$\begin{aligned}\widetilde{e}_i(\theta) &= e_i(\theta) + \xi_i \\ E\xi_i &= 0\end{aligned}$$

$$\begin{aligned}plim\left[\frac{1}{n}\sum_{i=1}^{n} Z_i^{\prime}\widetilde{e}_i(\theta)\right] &= plim\left[\frac{1}{n}\sum_{i=1}^{n} Z_i^{\prime}\left(e_i(\theta) + \xi_i\right)\right] \\ &= plim\left[\frac{1}{n}\sum_{i=1}^{n} Z_i^{\prime} e_i(\theta) + \frac{1}{n}\sum_{i=1}^{n} Z_i^{\prime}\xi_i\right] \\ &= plim\left[\frac{1}{n}\sum_{i=1}^{n} Z_i^{\prime} e_i(\theta)\right]\end{aligned}$$

MSM

- Loss in efficiency depends on properties of ξ_i

MSM

- Loss in efficiency depends on properties of ξ_i
- In worst case, $Var(\xi_i) = \frac{1}{R} Var(e_i)$

MSM

- Loss in efficiency depends on properties of ξ_i
- In worst case, $Var\left(\xi_i\right) = \frac{1}{R} Var\left(e_i\right)$
- If $e_i\left(\cdot\right)$ is monotone and one uses antithetic acceleration, the the variance due to simulation is $O\left(\frac{1}{n}\right)$

MSM

- Loss in efficiency depends on properties of ξ_i
- In worst case, $Var\left(\xi_i\right) = \frac{1}{R} Var\left(e_i\right)$
- If $e_i\left(\cdot\right)$ is monotone and one uses antithetic acceleration, the the variance due to simulation is $O\left(\frac{1}{n}\right)$
- Note that MSM relies only on using an unbiased simulator

MSL

$$L(\theta) = \sum_{i=1}^{n} \log L_i(\theta)$$

- In some problems (e.g., MNP), while we cannot evaluate $L_i(\theta)$ easily because it involves high-dimensional integrals, we can simulate it easily.

MSL

$$L(\theta) = \sum_{i=1}^{n} \log L_i(\theta)$$

- In some problems (e.g., MNP), while we cannot evaluate $L_i(\theta)$ easily because it involves high-dimensional integrals, we can simulate it easily.
- However, even if the simulator of $L_i(\theta)$ is unbiased, the simulator of $\log L_i(\theta)$ may have bad properties

MSL

$$L(\theta) = \sum_{i=1}^{n} \log L_i(\theta)$$

- In some problems (e.g., MNP), while we cannot evaluate $L_i(\theta)$ easily because it involves high-dimensional integrals, we can simulate it easily.
- However, even if the simulator of $L_i(\theta)$ is unbiased, the simulator of $\log L_i(\theta)$ may have bad properties(eg. MNP with frequency simulator)

MSL

$$L(\theta) = \sum_{i=1}^{n} \log L_i(\theta)$$

- In some problems (e.g., MNP), while we cannot evaluate $L_i(\theta)$ easily because it involves high-dimensional integrals, we can simulate it easily.
- However, even if the simulator of $L_i(\theta)$ is unbiased, the simulator of $\log L_i(\theta)$ may have bad properties(eg. MNP with frequency simulator)
- Consistency requires that $R \to \infty$ as $n \to \infty$

MSL

$$L(\theta) = \sum_{i=1}^{n} \log L_i(\theta)$$

- In some problems (e.g., MNP), while we cannot evaluate $L_i(\theta)$ easily because it involves high-dimensional integrals, we can simulate it easily.
- However, even if the simulator of $L_i(\theta)$ is unbiased, the simulator of $\log L_i(\theta)$ may have bad properties(eg. MNP with frequency simulator)
- Consistency requires that $R \rightarrow \infty$ as $n \rightarrow \infty$[compare to MSM]
- Borsch-Supan and Hajivassiliou showed, using Monte Carlo methods, that MSL has small asymptotic bias and, in fact, usually behaves better than MSM as long as one uses a simulator with good properties (eg, GHK)

MSL

$$L(\theta) = \sum_{i=1}^{n} \log L_i(\theta)$$

- In some problems (e.g., MNP), while we cannot evaluate $L_i(\theta)$ easily because it involves high-dimensional integrals, we can simulate it easily.
- However, even if the simulator of $L_i(\theta)$ is unbiased, the simulator of $\log L_i(\theta)$ may have bad properties(eg. MNP with frequency simulator)
- Consistency requires that $R \to \infty$ as $n \to \infty$[compare to MSM]
- Borsch-Supan and Hajivassiliou showed, using Monte Carlo methods, that MSL has small asymptotic bias and, in fact, usually behaves better than MSM as long as one uses a simulator with good properties (eg, GHK)
- Antithetic acceleration result

Other Topics

- MSS
- MCMC Methods
- Monte Carlo tests
- Parametric Bootstraps
- Distributions of Test Statistics
- Simulation inside of expected values (eg, value function approximation)