5.8 Summary

- Reduction of statistical error in MC quadrature possible through variance reduction
- Variety of possible methods
  - antithetic variables*
  - control variates*
  - matching moments*
  - importance sampling
  - stratification

*frequently used in finance

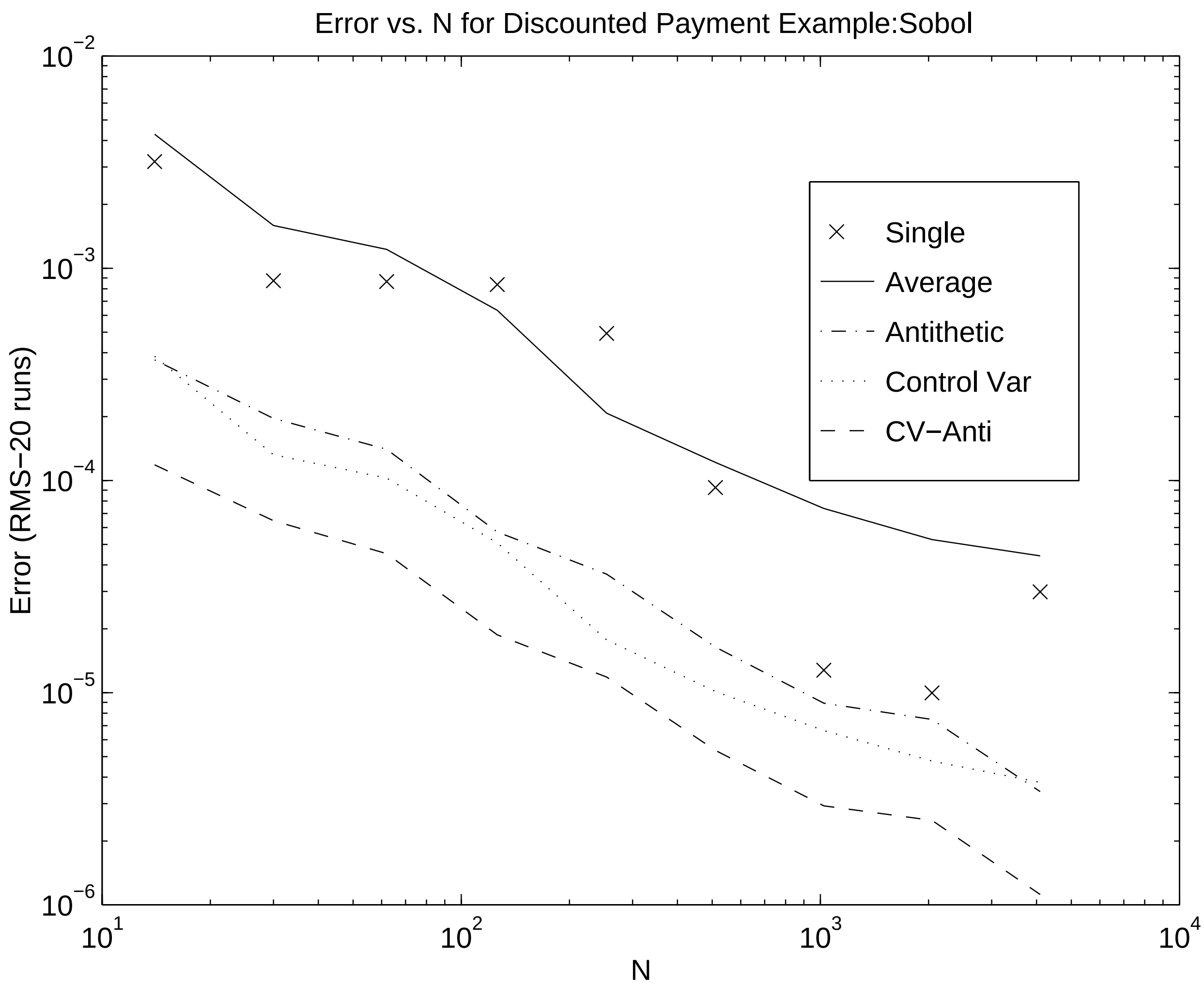

Figure 6: Discounted Cashflow, QMC

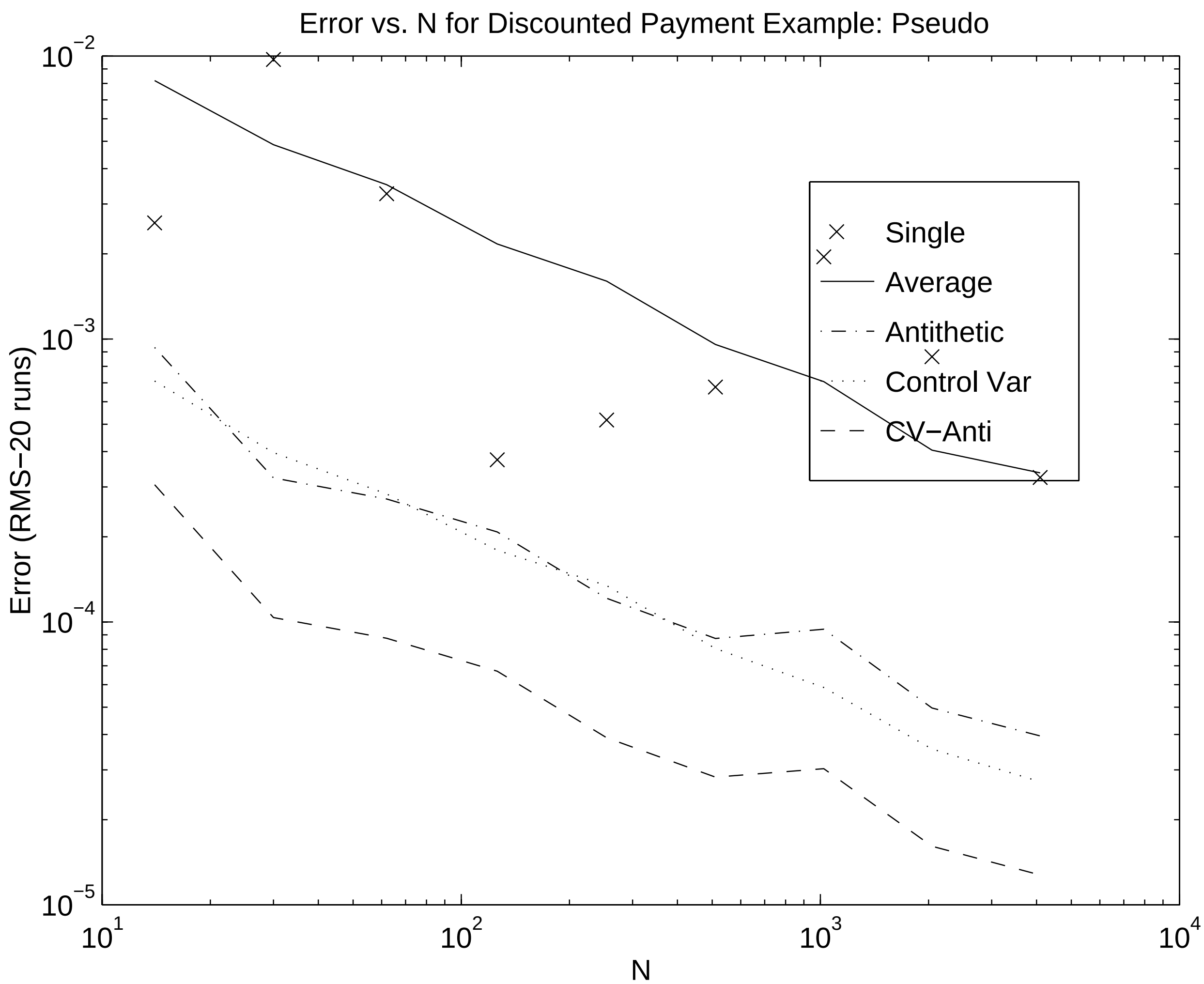

Figure 5: Discounted Cashflow, MC

Control variate

Approximate $(1+r_i)^{-1}$ by $1-r_i$ for $i \geq 1$.
Form control variate as

$$g = (1+r_0)^{-1}(1-r_1)(1-r_2)(1-r_3)$$

Since g consists of a sum of linear exponentials, its integral can be performed exactly; e.g.

$$\begin{aligned}\int_{-\infty}^{\infty} e^{\lambda x} e^{-x^2/2} dx &= e^{\lambda^2/2} \int_{-\infty}^{\infty} e^{-(x-\lambda)^2/2} dx \\ &= e^{\lambda^2/2} \int_{-\infty}^{\infty} e^{-y^2/2} dy \\ &= \sqrt{2\pi} e^{\lambda^2/2}\end{aligned}$$

Numerical Results

- Standard MC
- Quasi-MC (described in next section)
- Antithetic variables
- Control variates

Present value is

$$\begin{aligned} PV &= E[u] \\ &= (2\pi)^{-3/2} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u e^{-(x_1^2+x_2^2+x_3^2)/2} dx_1 dx_2 dx_3 \end{aligned}$$

Discount factor u and interest rates r_i are

$$\begin{aligned} u &= (1+r_0)^{-1}(1+r_1)^{-1}(1+r_2)^{-1}(1+r_3)^{-1} \\ r_1 &= r_0 e^{\sigma x_1 - \sigma^2/2} \\ r_2 &= r_1 e^{\sigma x_2 - \sigma^2/2} \\ r_3 &= r_2 e^{\sigma x_3 - \sigma^2/2} \end{aligned}$$

Parameters

- initial interest rate $r_0 = .10 = 10\%$
- size of interest rate fluctuations $\sigma = .1$ (1% per year)

Evaluate PV by sampling x_i from $N(0,1)$ distribution using transformation method. Apply antithetic variables and control variates.

5.7 Example – Discounted Payment

Present value of discounted payment.

- Payment of \$1 after $M = 4$ years
- Year i
  - interest rate r_i
  - annual discount factor $(1 + r_i)^{-1}$
- log normal interest rate model

$$r_i = r_{i-1} e^{\sigma x_i - \sigma^2/2}$$

- x_i is $N(0, 1)$
- Simple path dependent security

Effectiveness of importance sampling

- If f/p is nearly constant, then

$$\sigma_p << \sigma$$

.

- Sample from $p(x)$ using acceptance-rejection, if necessary.
- Use to emphasize rare events
  - risk
  - far out of the money options

5.6 Importance Sampling

Rewrite integral using density p.

- Integral

$$I(f) = \int f(x)dx = \int \frac{f(x)}{p(x)} p(x)dx.$$

- Monte Carlo estimate is

$$I_N(f) = \frac{1}{N} \sum_{n=1}^{N} \frac{f(x_n)}{p(x_n)}.$$

Error $\epsilon_N(f) = I(f) - I_N(f)$

- Error size

$$\epsilon_N(f) \approx \sigma_p N^{-1/2}$$

- Variance

$$\sigma_p = \int \left(\frac{f(x)}{p(x)} - I \right)^2 p(x)dx$$

Stratification always lowers integration error if distribution of points is balanced.

- balance condition: for all k

$$\bar{p}_k / N_k = 1/N$$

- number of points in set Ω_k is proportional to its weighted size $\bar{p}_k$
- resulting error for stratified quadrature

$$\begin{aligned} \epsilon_N &\approx N^{-1/2} \sigma_s \\ \sigma_s^2 &= \sum_{k=1}^{M} \sigma^{(k)2} \end{aligned}$$

- Variance reduction

$$\sigma_s \leq \sigma$$

- Resulting error reduction

$$\epsilon_{sN} \leq \epsilon_N$$

- Better choice - put more points where f has largest variation

Resulting integration error

- Total error

$$\begin{aligned} \epsilon_{sN}(f) &= I(f) - I_N(f) \\ &= \sum_{k=1}^{M} \epsilon_{N_k}^{(k)}(f) \end{aligned}$$

- Error components

$$\begin{aligned} \epsilon_{N_k}^{(k)}(f) &\approx N_k^{-1/2} \bar{p}_k \left(\int_{\Omega_k} (f(x) - \bar{f}_k)^2 p^{(k)}(x) dx \right)^{1/2} \\ &= (\bar{p}_k / N_k)^{1/2} \sigma^{(k)} \end{aligned}$$

- Variances

$$\begin{aligned} \sigma^{(k)} &= (\bar{p}_k)^{1/2} \left(\int_{\Omega_k} (f(x) - \bar{f}_k)^2 p^{(k)}(x) dx \right)^{1/2} \\ &= \left(\int_{\Omega_k} (f(x) - \bar{f}_k)^2 p(x) dx \right)^{1/2} \end{aligned}$$

- Averages

$$\bar{f}_k = \int_{\Omega_k} f(x) p(x) dx / \bar{p}_k.$$

General Formulation of Stratification

- Split integration region Ω into M pieces Ω_k with

$$\Omega = \cup_{k=1}^{M} \Omega_k$$

- Take N_k random variables in each piece Ω_k with

$$\sum_{k=1}^{M} N_k = N$$

- $x_n^{(k)}$ distributed with density $p^{(k)}(x)$ in Ω_k
  - $p^{(k)}(x) = p(x)/\bar{p}_k$
  - $\bar{p}_k = \int_{\Omega_k} p(x)dx$
  - note: $\int_{\Omega_k} p^{(k)}(x)dx = 1$
- Stratified quadrature formula is sum over k

$$I_N(f) = \sum_{k=1}^{M} \frac{1}{N_k} \sum_{n=1}^{N_k} f(x_n^{(k)})$$

MC Integration Error Using Stratification

- MC quadrature error

$$\begin{aligned}\epsilon &\approx N^{-1/2}\sigma_s \\ \sigma_s^2 &= \int \left(f(x) - \bar{f}(x)\right)^2 dx \\ &= \sum_{k=1}^{M} \int_{\Omega_k} \left(f(x) - \bar{f}_k\right)^2 dx\end{aligned}$$

Stratified always beats non-stratified

$$\sigma_s \leq \sigma$$

Proof: For each k, $c = \bar{f}_k$ is minimizer for

$$\int_{\Omega_k} \left(f(x) - c\right)^2 dx$$

In particular,

$$\int_{\Omega_k} \left(f(x) - \bar{f}_k\right)^2 dx \leq \int_{\Omega_k} \left(f(x) - \bar{f}\right)^2 dx$$

Sum this to get

$$\begin{aligned}\sigma_s^2 &= \sum_{k=1}^{M} \int_{\Omega_k} \left(f(x) - \bar{f}_k\right)^2 dx \\ &\leq \sum_{k=1}^{M} \int_{\Omega_k} \left(f(x) - \bar{f}\right)^2 dx \\ &= \sigma^2\end{aligned}$$

5.5 Stratification

Combination of grid and random variables.

Simplest case: Stratification based on regular grid with uniform density

- Split integration region $\Omega = [0, 1]$ into M pieces Ω_k

$$\Omega_k = \left[\frac{(k-1)}{M}, \frac{k}{M}\right]$$

- $|\Omega_k| = 1/M$
- For each k, sample $N_k = N/M$ points $\{x_i^{(k)}\}$ uniformly distributed in Ω_k
- Averages

$$\bar{f}(x) = \bar{f}_k = M \int_{\Omega_k} f(x)dx \quad \text{for } x \in \Omega_k.$$

- Quadrature formula

$$I_N = N^{-1} \sum_{k=1}^{M} \sum_{i=1}^{N/M} f(x_i^{(k)})$$

Transformation of sample points

- Match first moment

$$y_n = (x_n - \mu_1) + m_1$$

  – satisfies $\frac{1}{N}\Sigma\, y_n = m_1$

- Match first two moments

$$\begin{aligned} y_n &= (x_n - \mu_1)/c + m_1 \\ c &= \sqrt{\frac{m_2 - m_1^2}{\mu_2 - \mu_1^2}} \end{aligned}$$

  – $\frac{1}{N}\Sigma\, y_n = m_1$
  – $\frac{1}{N}\Sigma\, y_n^2 = m_2$

Caution: Sample points no longer independent

- CLT not applicable
- Error estimates less straightforward

5.4 Matching Moments Method

Monte Carlo integration error partly due to statistical sampling error

- Distribution of $\{x_n\}_{n=1}^N$ not exactly $p(x)$
- E.g. $\mu_1 \neq m_1$, $\mu_2 \neq m_2$
- First and second moments of p

$$\begin{aligned} m_1 &= \int x p(x) dx \\ m_2 &= \int x^2 p(x) dx \end{aligned}$$

- First and second moments of sample $\{x_n\}_{n=1}^N$

$$\begin{aligned} \mu_1 &= N^{-1} \sum_{n=1}^{N} x_n \\ \mu_2 &= N^{-1} \sum_{n=1}^{N} x_n^2 \end{aligned}$$

- Partial correction: make moments exact

Optimal Use of Control Variate

Introduce optimal multiplier λ for control variate g

- Integral

$$\int f(x)dx = \int (f(x) - \lambda g(x))\, dx + \lambda \int g(x)dx$$

- Error in first integral is proportional to variance

$$\sigma^2_{f-\lambda g} = \int \left(\tilde{f}(x) - \lambda \tilde{g}(x)\right)^2 dx$$

in which

$$\begin{aligned} \tilde{f}(x) &= f(x) - I[f] \\ \tilde{g}(x) &= g(x) - I[g] \end{aligned}$$

- Optimal value of λ found by minimizing $\sigma^2_{f-\lambda g}$ to obtain

$$\begin{aligned} \lambda &= E[\tilde{f}\tilde{g}]/E[\tilde{g}^2] \\ &= \left(\int \tilde{f}\tilde{g}dx\right) / \left(\int \tilde{g}^2 dx\right) \end{aligned}$$

5.3 Control Variates

Use integrand g, which is similar to f.

- Integral

$$\int f(x)dx = \int (f(x) - g(x))\, dx + \int g(x)dx$$

- Known integral $I(g) = \int g(x)dx$
- Monte Carlo quadrature formula

$$I_n(f) = \frac{1}{N} \sum_{n=1}^{N} (f(x_n) - g(x_n)) + I(g)$$

Integration error $\epsilon_N(f) = I(f) - I_N(f)$

- Size

$$\epsilon_N(f) \approx \sigma_{f-g} N^{-1/2}$$

- Variance

$$\sigma_{f-g}^2 = \int \left(f(x) - g(x) - (\bar{f} - \bar{g})\right)^2 dx.$$

Effective if

$$\sigma_{f-g} << \sigma_f$$

5.2 Antithetic Variables

For each sample value x also use the value $-x$.

Motivation: $E[f(x)]$ with x from $N(0, \sigma^2)$

- $x = \sigma\hat{x}$
- Taylor expansion of $f = f(\sigma\hat{x})$ (for small σ) as

$$f = f(0) + f'(0)\sigma\hat{x} + O(\sigma^2).$$

- In $I = E[f]$ linear terms have average 0
- Linear terms cancel exactly with antithetic variables

5.1 Motivation

Integration error ϵ in Monte Carlo with N samples are related by

$$\begin{aligned} \epsilon &= O(\sigma N^{-1/2}) \\ N &= O(\sigma/\epsilon)^2 \end{aligned}$$

Two options for acceleration (error reduction)

- Variance reduction
  - transform the integrand to reduce the variance σ
- Modified statistics
  - Replace random variables (e.g. by *quasi-random variables*)

Caution: Acceleration method may require extra computational time, which must be balanced against reduced N.

5. Variance Reduction

5.1 Motivation

5.2 Antithetic Variables

5.3 Control Variates

5.4 Matching Moments

5.5 Stratification

5.6 Importance Sampling

5.7 Example – Discounted Payment

5.8 Summary

4.5 Summary

- European options, Greeks, basket options easily calculated by Monte Carlo
  - accuracy $= O(N^{-1/2})$, cpu time $= O(N)$
  - slow compared to PDE or tree methods for single security
  - faster for multiple securities

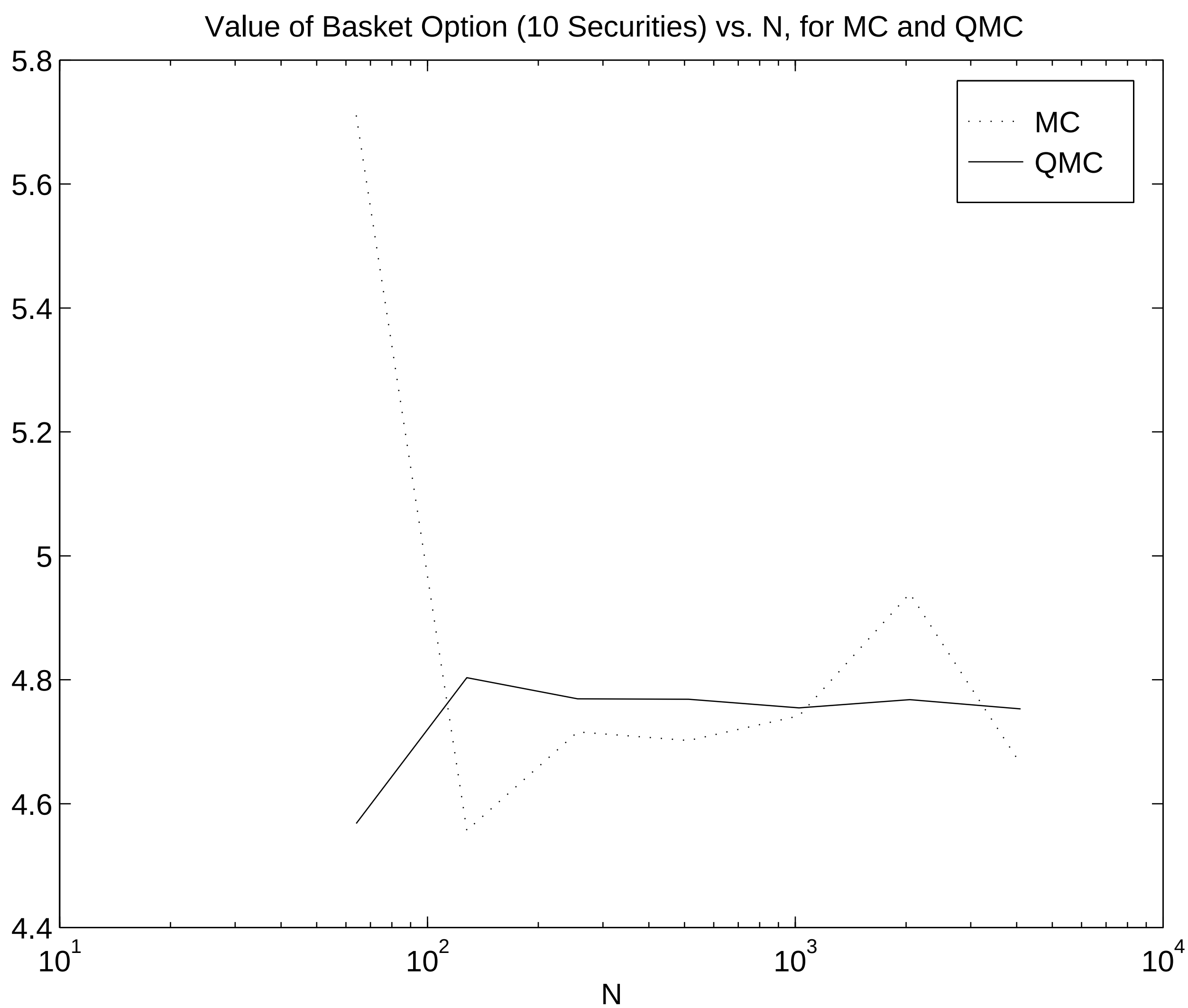

Figure 4: Basket Option with 10 Securities

4.4 Basket Options

Simple basket option

- European call for arithmetic average of M securities
- Securities are correlated

$$dS_i = rS_i dt + \sum_{j=1}^{M} \sigma_{ij} S_j db_j(t)$$

- $b_j(t)$ independent standard Brownian motions
- Pricing formulas

$$\begin{aligned} S_i(t) &= S_i(0) \exp\left(rt + \sum_{j=1}^{M} (\sigma_{ij} \sqrt{t} \nu_j - \frac{1}{2} \sigma_{ij}^2 t) \right) \\ S_{av} &= \frac{1}{M} \sum_{i=1}^{M} S_i(t) \\ PV_C &= e^{-rt} E\left(\max\left(0, S_{av} - K\right)\right) \end{aligned}$$

- ν_j are independent $N(0,1)$ random variables
- M-dimensional integral
- Difficult for pde or lattice method if $M \geq 3$
- Monte Carlo performs well independent of dimension
- Computational example $M = 10$

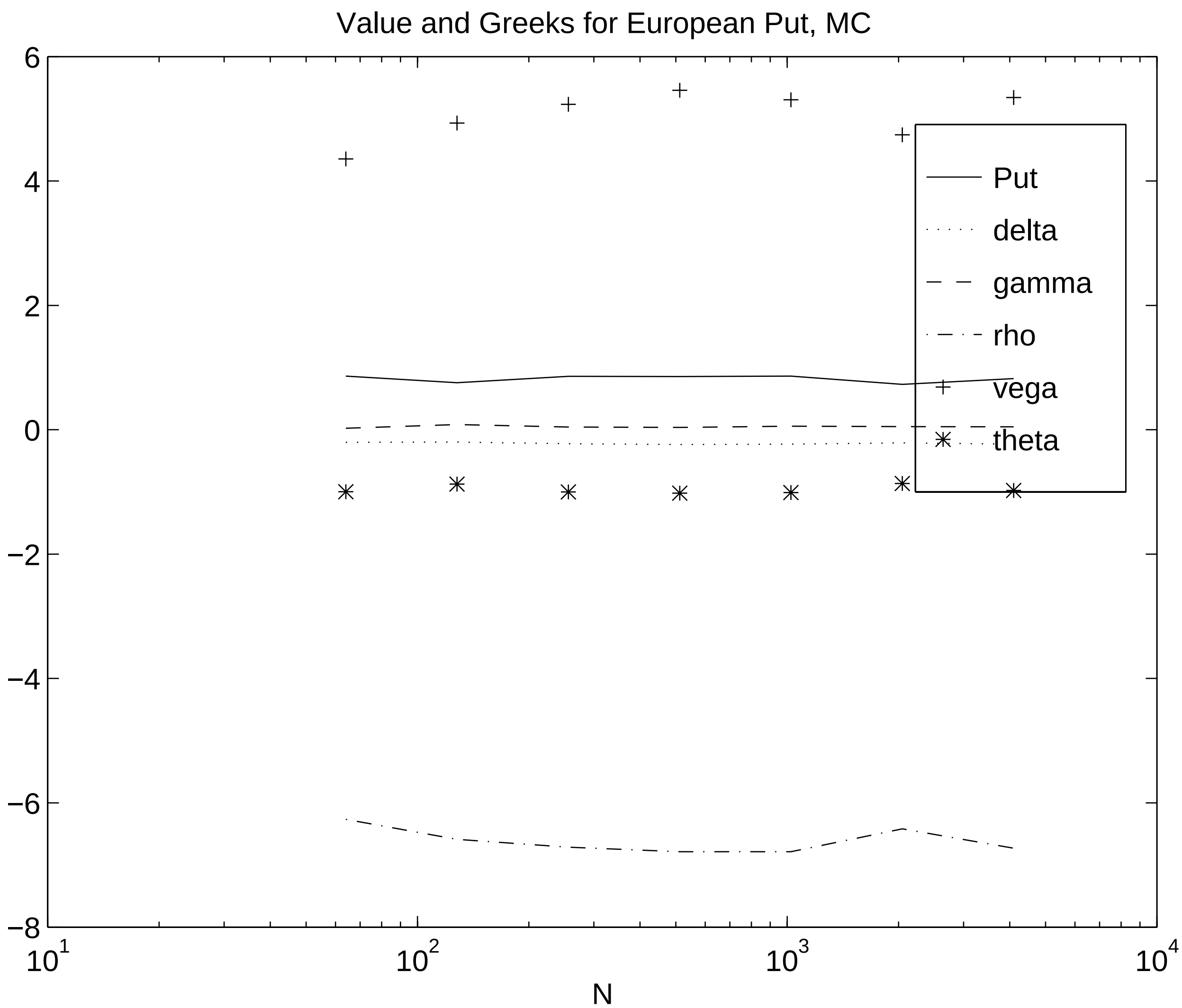

Figure 3: Value and Greeks for Put using MC

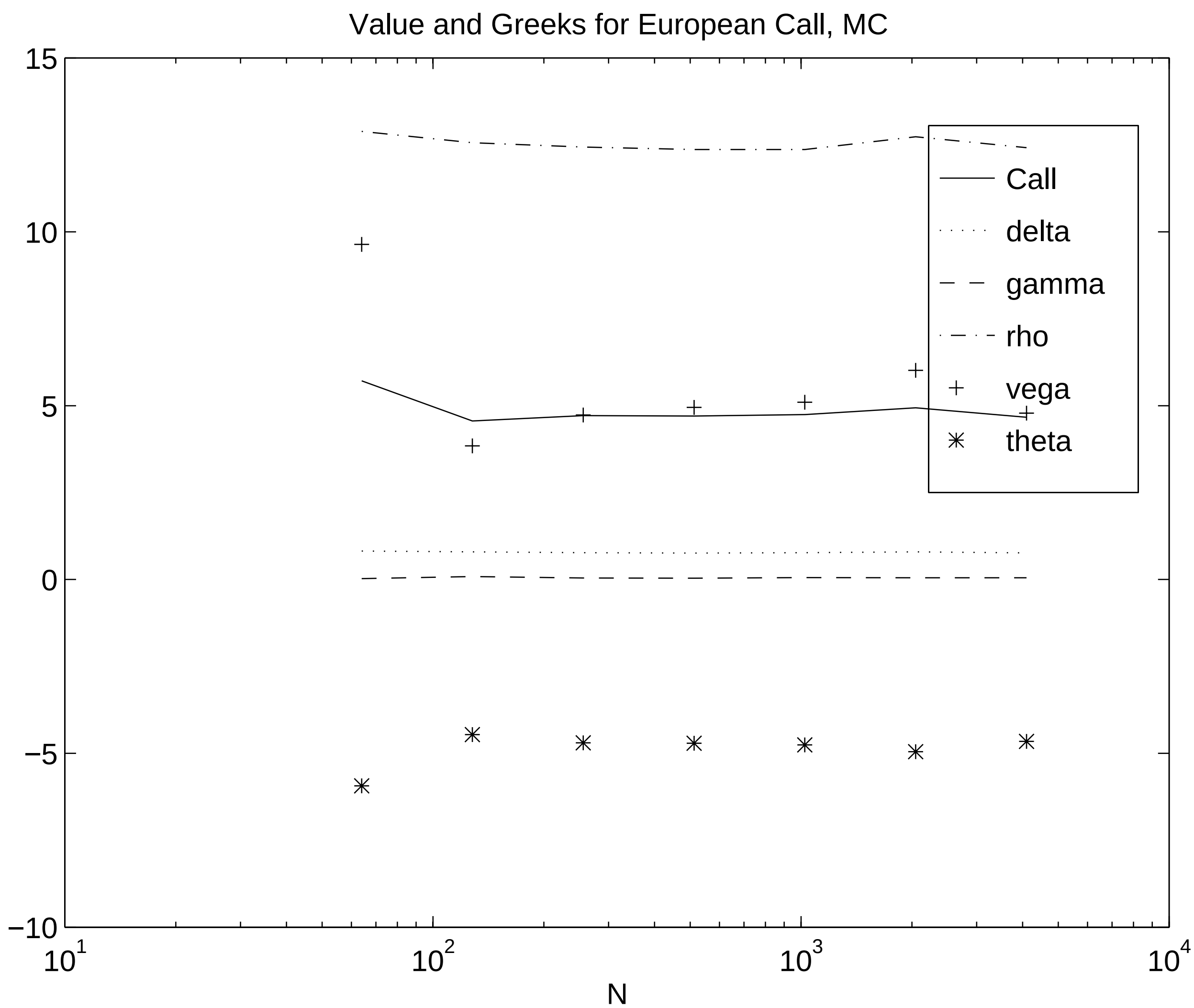

Figure 2: Value and Greeks for Call using MC

4.3 Computational Results

Parameters

- $S_0 = 42$
- $K = 40$
- $r = .1$
- $\sigma = .2$
- $t = .5$
- Call is in-the-money
- Put is out-of-the-money

4.2 Monte Carlo Calculation of the Greeks

- Exact formulas possible for vanilla options, but not for exotics
- Approximate derivatives by finite difference; e.g. for $\Delta_C = \partial PV_C / \partial S_0$ and $\Gamma_C = \partial^2 PV_C / \partial S_0^2$

$$\Delta_C \approx \frac{PV_C(S_0 + \delta) - PV_C(S_0 - \delta)}{2\delta}$$

$$\Gamma_C \approx \frac{PV_C(S_0 + \delta) + PV_C(S_0 - \delta) - 2PV_C(S_0)}{\delta^2}$$

- Use same random numbers for $PV_C(S_0)$, $PV_C(S_0 + \delta)$ and $PV_C(S_0 - \delta)$.
- Minimizes statistical error

Notation

- μ = rate of return on stock
- r = risk free rate of return
- σ = volatility
- $b(t)$ = standard Brownian motion
- S_0 = current price
- K = strike price
- ν is $N(0, 1)$

4.1 European Options

Stock price $S(t)$

- Real process $dS = \mu S dt + \sigma S db$
- Risk-neutral process $dS = rS dt + \sigma S db$
  - use risk-neutral process in subsequent formulas
  - from Ito calculus

$$S(t) = S_0 \exp((r - \sigma^2/2)t + \sigma\sqrt{t}\nu)$$

Call and Put Values

- Call Value PV_C

$$PV_C = e^{-rt}E(\max(0, S(t) - K))$$

- Put Value PV_P

$$PV_P = e^{-rt}E(\max(0, K - S(t)))$$

- Equivalent to Black-Scholes formulas
- Evaluate by sampling ν, using formula for $S(t)$

4. Applications: European Options, the Greeks

3.4 Summary

- Monte Carlo quadrature is straightforward and robust
- Empirical estimates of accuracy as part of computation
- Determine necessary N for desired accuracy

Variation of N

For desired accuracy level $\bar{\varepsilon}$

- Perform series of M computations at some initial value N_0
- Calculate RMS error $\tilde{\varepsilon}_{N_0}$
- Obtain desired accuracy by using $N = N_0(\tilde{\varepsilon}_{N_0}/\bar{\varepsilon})^2$

Note

- At this point already computed with $N_1 = MN_0$
- Easy to increment values as N increases
- Use independent values of x_i to assess accuracy vs. N
- For final value at largest value of N add on additional values of f_i using formula

$$I_N = (n/N)I_n + N^{-1}\sum_{i=n+1}^{N} f_i$$

3.3 Computation and Interpretation of Results

Monte Carlo method illustrated by comparing value I_N at different values of N

Accuracy of results is assessed, as described in Section 1.3

- Perform M independent calculations of $I_N^{(j)}$ with $M \approx 20$
- Average value $\bar{I}_N = M^{-1} \sum I_N^{(j)}$
- Empirical RMS error $\tilde{\varepsilon}_N$

$$\tilde{\varepsilon}_N = \left(M^{-1} \sum (I_N^{(j)} - \bar{I}_N)^2 \right)^{1/2}$$

Plot $\tilde{\varepsilon}_N$ vs. N to see $1/\sqrt{N}$ behavior

3.2 Monte Carlo Code

choose $N = \#samples$

generate uniform random variables y_i

transform to $N(0,1)$ variables x_i

evaluate and sum $f_i = f(x_i) = x_i^2$

divide by N

result $I_N = N^{-1} \sum f_i$

3.1 Example – Gaussian Integral

Consider the following Gaussian integral

$$\begin{aligned} I &= (2\pi)^{-1/2} \int_{-\infty}^{\infty} x^2 e^{-x^2/2} dx \\ &= E[x^2] \end{aligned}$$

- Expectation is over $N(0, 1)$ random variable x
- Exact value $I = 1$
- Compute using Monte Carlo method

MC quadrature formula is

$$I_N = N^{-1} \sum_{i=1}^{N} x_i^2$$

3. Implementation of Monte Carlo Computational Method

3.1 Example – Gaussian Integral

3.2 Monte Carlo Code

3.3 Computation and Interpretation of Results

3.4 Summary

2.7 Summary

- Pseudo-random numbers generated by standard methods
  - uniform distribution on $[0, 1]$
  - warning: very good but not perfect
- Sampling of nonuniform distribution from uniform distribution
  - general methods: transformation and acceptance-rejection
  - special methods: Box-Muller and others

Effectiveness of Acceptance-Rejection:

- Advantage: General
- Normalization of p not required
- Advantage: Requires no inversion of P
- Disadvantage: May be inefficient, requiring many trials before acceptance

Acceptance-Rejection: Partial Derivation

Since $p < q$, then

$$\begin{aligned} \int_0^1 \chi(\frac{p(x)}{q(x)} > y) dy &= \frac{p(x)}{q(x)} \\ &= p(x)(\hat{q}(x) I[q])^{-1} \end{aligned}$$

So

$$\begin{aligned} \int f(x) p(x) dx &= \int \int_0^1 f(x)\ \chi(\frac{p(x)}{q(x)} > y)\ \hat{q}(x)\ dy\ dx\ I[q] \\ &\approx N'^{-1} \sum_{p(x'_n)/q(x'_n) > y_n} f(x'_n)\ I[q] \\ &\approx N^{-1} \sum_{\text{accepted pts}} f(x_n) \end{aligned}$$

in which

$$\begin{aligned} N' &= \text{total number of trial points} \\ N &= \text{total number of accepted points} \\ &\approx N'/I[q] \end{aligned}$$

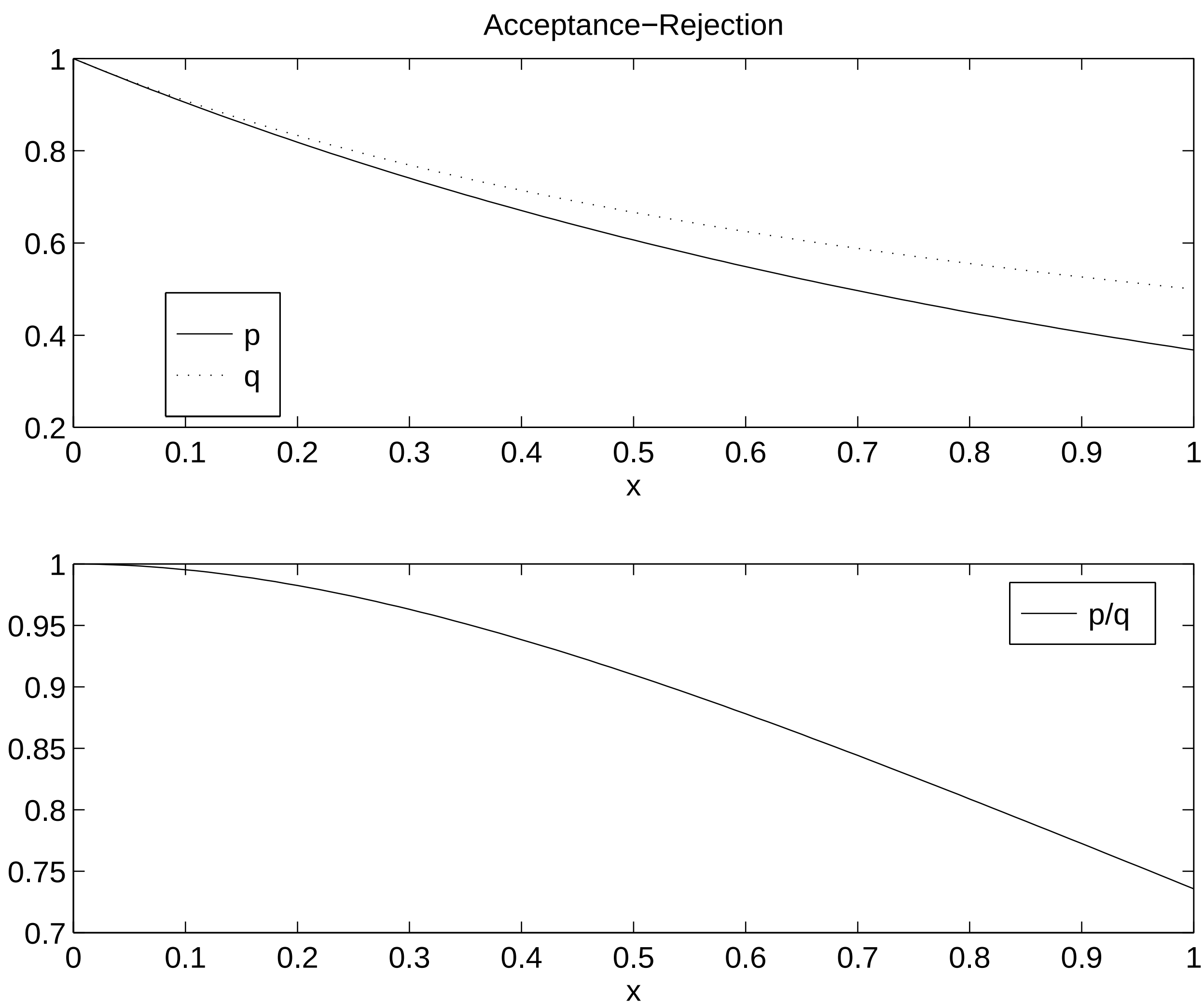

Figure 1: Typical choice of p, q and p/q. Accept for $y < p/q$ and reject for $y > p/q$.

2.6 Acceptance-Rejection Method

General way of producing random variables of given density $p(x)$

- Choose $q(x) \geq p(x)$
- Probability density

$$\begin{aligned} \hat{q}(x) &= q(x)/I[q] \\ I[q] &= \int q(x')dx' \end{aligned}$$

Acceptance-Rejection Procedure

- Two random variables, x', y
  - trial variable x' with density $\hat{q}(x')$
  - decision variable y uniform $(0 < y < 1)$
- *accept* if $0 < y < p(x')/q(x')$
- *reject* if $p(x')/q(x') < y < 1$
- Repeat until a value x' is accepted
- Once x' is accepted, take

$$x = x'.$$

Box-Muller is based on the following observation:

- Change variables $(x_1, x_2) = (r\cos\theta, r\sin\theta)$ to polar coordinates (r, θ)

$$dx_1 dx_2 = r dr d\theta$$

- Integration element

$$(2\pi)^{-1} e^{-(x_1^2+x_2^2)/2} dx_1 dx_2 = (2\pi)^{-1} e^{-r^2/2} r dr d\theta$$

- Angular variable $\theta/(2\pi)$ is uniformly distributed.
- Variable r is easily sampled
  - density $re^{-r^2/2}$
  - distribution

$$\begin{aligned} P(r) &= \int_0^r e^{-r'^2/2} r' dr' = 1 - e^{-r^2/2} \\ P^{-1}(y) &= \sqrt{-2\log(1-y)} \end{aligned}$$

- Resulting transform

$$(y_1, y_2) \to (r, \theta) \to (x_1, x_2)$$

2.5 Box-Muller Method

- Direct way of generating normal random variables without inverting the error function.
- y_1, y_2 two uniform variables
- Obtain two normal variables x_1, x_2:

$$
\begin{aligned}
x_1 &= \sqrt{-2\log(y_1)}\cos(2\pi y_2) \\
x_2 &= \sqrt{-2\log(y_1)}\sin(2\pi y_2)
\end{aligned}
$$

2.4 Gaussian (Normal) Random Variables

Normal random variable: density p and distribution P

$$
\begin{aligned}
p(x) &= (2\pi)^{-1/2} e^{-x^2/2} \\
P(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt \\
&= \frac{1}{2} + \frac{1}{2} \mathrm{erf}(x/\sqrt{2})
\end{aligned}
$$

in which

- Factor $(2\pi)^{-1/2}$ is a normalization
- $\int_{-\infty}^{\infty} p(x)dx = 1$
- Error function erf defined by

$$\mathrm{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt.$$

Sample normal variable x, using uniform variable y, by

$$y = P(x) = \frac{1}{2} + \frac{1}{2}\ \mathrm{erf}(x/\sqrt{2})$$

i.e.,

$$x = \sqrt{2}\ \mathrm{erf}^{-1}(2y - 1).$$

2.3 Transformation Method

Goal: Produce random variable x with density $p(x)$.

Method: Write $x = X(y)$ with y a uniform variable.

Define probability distribution function

$$P(x) = \int_{-\infty}^{x} p(x')dx'$$

Determination of the mapping $X(y)$:

$$\begin{aligned}
E_p[f(x)] &= E_{unif}[f(X(y))] \text{ for any } f \\
&\Longrightarrow \int f(x)p(x)dx = \int f(X(y))dy \\
&\qquad\qquad = \int f(x)(dy/dx)dx \text{ by change of variables} \\
&\Longrightarrow p(x) = dy/dx \\
&\Longrightarrow \int^{X(y)} p(x)dx = dy \\
&\Longrightarrow P(X(y)) = y \\
&\Longrightarrow X(y) = P^{-1}(y)
\end{aligned}$$

Summary

- y (uniform) $\rightarrow$ x (density $p(x)$)
- $x = P^{-1}(y)$, inverse function of distribution P

2.2 Non-Uniform Density

- Standard random number generators produce uniformly distributed variables.
- Non-uniform variables sampled through transformation of uniform variable.
- For a non-uniform random variable with density $p(x)$

$$\begin{aligned} E[f] &= I[f] = \int f(x)p(x)dx \\ \epsilon_N(f) &= I(f) - I_N(f) \\ &= \int f(x)p(x)dx - \frac{1}{N}\sum_{n=1}^{N} f(x_n) \end{aligned}$$

CLT results:

- $\epsilon_N(f) \approx N^{-1/2}\sigma\nu$
- ν is $N(0,1)$
- $\sigma^2 = \int (f - \bar{f})^2 p(x)dx$

2.1 Random Number Generators

Numbers generated by computers are not random, but pseudo-random

- Pseudo-random sequences are made to have many properties of random numbers
  - well-developed subject
  - occasional problems still occur, mostly with very long sequences ($N \geq 10^9$)
  - linear congruential methods predominate
- Series of reliable methods in Numerical Recipes, 2nd Ed.
  - ran0, ran1, ran2, ran3, ran4
  - ran1 recommended for $N < 10^8$
- Very poor methods in Numerical Recipes, 1st Ed.
  - DO NOT USE RAN0, RAN1, RAN2, RAN3

2. Generating and Sampling Random Variables

2.1 Random Number Generators

2.2 Non-Uniform Density

2.3 Transformation Method

2.4 Gaussian (Normal) Random Variables

2.5 Box-Muller Method

2.6 Acceptance-Rejection Method

2.7 Summary

1.7 Summary

- Monte Carlo quadrature based on probabilistic representation of integrals
- Error ε is of size $\sigma N^{-1/2}$ for N samples and variance σ
- Empirical estimate of σ provides guide to necessary sample size N for desired accuracy ε with confidence c
- Monte Carlo also applicable to simulation and optimization

1.6 When Should Monte Carlo Be Used?

- High dimension
  - e.g. basket options with 3 or more securities
- Finite difference or tree methods unavailable
  - e.g. path-dependent securities
- Simulation
  - e.g. value at risk

How can irregular set of points beat a grid?

- Fourier interpretation (periodic functions)
  - grid of spacing $1/n$ gives
    - 0% accuracy on wavenumber $k = mn$ (for m a nonzero integer)
    - 100% accuracy on wavenumbers $k \neq mn$
  - random array gives partial accuracy for all k
- Tree interpretation
  - in regular grid, variation of only one component at a time; i.e.

$$(0, 0, \ldots, 0, 0) \rightarrow (0, 0, \ldots, 0, 1/n)$$

  - inefficient use of unvaried components
  - random array, all components varied in each point

1.5 Comparison to Grid-Based Methods

Comparison to grid-based integration methods (e.g. Simpson's rule)

- Convergence rate for grid based quadrature: $O(N^{-k/d})$ for order k method in dimension d
  - MC beats grid in high dimension d ($d \geq 3$ or 4),
$$k/d < 1/2$$
- Impossible to lay down grid in high dimension
  - requires at least 2^d points
- Impossible to refine grids in high dimension
  - requires increasing # points by factor 2^d points

1.4 Simulation and Optimization

Quadrature for valuation and risk is the focus of this course.
But there are additional uses for Monte Carlo:

Monte Carlo provides a direct and robust simulation of a financial process.

- E.g. interest rate paths
- Provides picture of future scenarios
- Represents full statistics
- Risk calculations

Optimization performed using Monte Carlo

- Search state space for largest value of payoff function $f(x)$
- E.g. optimal cash allocations

Use of CLT in Applications

Problem: Exact value is unknown (that's the desired value), so errors and the variance cannot be determined

Solution: Determine empirical error and variance

- Perform M computations using independent points x_i for $1 \leq i \leq MN$
- Obtain values $I_N^{(j)}$ for $1 \leq j \leq M$,
- Empirical RMS error is $\tilde{\varepsilon}_N$ given by

$$\tilde{\varepsilon}_N = \left(M^{-1} \sum_{j=1}^{M} (I_N^{(j)} - \bar{I}_N)^2 \right)^{1/2}$$

  in which

$$\bar{I}_N = M^{-1} \sum_{j=1}^{M} I_N^{(j)}$$

- Empirical variance is $\tilde{\sigma}$ given by

$$\tilde{\sigma} = N^{1/2} \tilde{\varepsilon}_N$$

Converse of CLT:

To insure

- Error of size ϵ
- With confidence level c

Require

- Number of sample points N given by

$$N = \epsilon^{-2}\sigma^2 s(c)$$

- s is confidence function for a normal variable.

$$\begin{aligned} c &= \int_{-s(c)}^{s(c)} e^{-x^2/2} dx / \sqrt{2\pi} \\ &= erf(s(c)/\sqrt{2}) \end{aligned}$$

Values of confidence function

confidence level c	0	.68	.80	.90	.95
multiplier $s(c)$	0	1.0	1.28	1.65	1.96

Partial Derivation of CLT

First define $\xi_i = \sigma^{-1}(f(x_i) - \bar{f})$ for x_i uniformly distributed. Then

- $E[\xi_i] = 0$
- $E[\xi_i^2] = 1$ since

$$\begin{aligned} E[\xi_i^2] &= \int \sigma^{-2}(f(x_i) - \bar{f})^2 dx \\ &= 1 \end{aligned}$$

- $E[\xi_i \xi_j] = 0$ if $i \neq j$ since independent

Variance of sum:

- Define $S_N = (1/N)\sum_1^N \xi_i = \sigma^{-1}\varepsilon_N$
- Calculate

$$\begin{aligned} E[S_N^2]^{1/2} &= E[N^{-2}(\sum_{i=1}^{N} \xi_i)^2]^{1/2} \\ &= N^{-1}\left\{E[\sum_{i=1}^{N} \xi_i^2] + E[\sum_{i=1}^{N}\sum_{j\neq i} \xi_i \xi_j]\right\}^{1/2} \\ &= N^{-1}\left\{\sum_{i=1}^{N} 1 + 0\right\}^{1/2} \\ &= N^{-1/2} \end{aligned}$$

- Therefore $RMSE_x = O(\sigma N^{-1/2})$

1.3 Central Limit Theorem (CLT)

CLT describes the size and statistical properties of Monte Carlo integration error.

- Integration error

$$\epsilon_N(f) = I(f) - I_N(f).$$

- CLT: For N large,

$$\epsilon_N(f) \approx \sigma N^{-1/2} \nu$$

- ν is $N(0,1)$ random variable
- $\sigma = \sigma(f)$ is the variance of f

$$\sigma(f) = \left(\int_{I^d} (f(x) - I(f))^2 dx \right)^{1/2}$$

- Precise statement

$$\lim_{N \to \infty} Prob(a < \frac{\sqrt{N}}{\sigma} \epsilon_N < b) = Prob(a < \nu < b)$$

Monte Carlo Quadrature Formula

- Sequence $\{x_n\}$ sampled from uniform distribution
- Empirical approximation to the expectation is

$$I_N[f] = \frac{1}{N} \sum_{n=1}^{N} f(x_n)$$

- Unbiased:

$$E[I_N[f]] = I[f]$$

- Convergent;

$$\lim_{N\to\infty} I_N[f] \to I[f]$$

In general define

- Error $\epsilon_N(f) = I(f) - I_N(f)$
- Bias $= E[\epsilon_N(f)]$
- RMSE="root mean square error" $= E[\epsilon_N(f)^2]^{1/2}$

1.2 Expectation and Integration

The integral of a function $f(x)$ can be expressed as the average or *expectation*. For one dimensional unit interval

- Integral $I[f] = \int_0^1 f(x)dx = \bar{f}$
- Average $E[f(x)] = \int_0^1 f(x)dx$ for a uniform random variable

For unit cube $I^d = [0,1]^d$ in d dimensions,

- $I[f] = E[f(x)] = \int_{I^d} f(x)dx$

Examples:

- Uniform: For $0 < x < 1$

$$p(x) = 1.$$

  Any number in unit interval is equally likely.

- Gaussian (a.k.a. normal or $N(0,1)$)

$$p(x) = (2\pi)^{-1/2} e^{-x^2/2}.$$

- Gaussian with mean m and variance σ^2 ($N(m,\sigma)$)

$$p(x) = (2\sigma^2\pi)^{-1/2} e^{-(x-m)^2/2\sigma^2}.$$

- Note: These are all continuous random variables

Discrete Random Variables

- Binomial

$$x = \begin{cases} 1 \\ -1 \end{cases} \text{with probability} \begin{cases} 1/2 \\ 1/2 \end{cases}$$

  $p(x) = 1/2\delta(x+1) + 1/2\delta(x-1).$

1.1 Random variables

A *random variable* X with a density function $p(x)$ takes on value x with relative probability $p(x)$.
Properties:

- $P(a < X < b) = \int_a^b p(x)dx$
- $0 \leq p(x)$
- $\int p(x)dx = 1$
- Mean $\bar{x} = \int xp(x)dx$
- Variance $\sigma^2 = \int (x - \bar{x})^2 p(x)dx$

1. Monte Carlo Integration

Outline

1. Monte Carlo Integration
2. Generating and Sampling Random Variables
3. Implementation of Monte Carlo Computation
4. Applications: European Options, Greeks
5. Variance Reduction

Monte Carlo Methods

for

Finance

Russel E. Caflisch

Mathematics Department

UCLA