Efficient Risk Estimation via Nested Sequential Simulation

Mark Broadie, Columbia University

Joint work with Yiping Du and Ciamac Moallemi

Bachelier Finance Society

June 24, 2010

Risk Measurement

Security positions today

- Hundreds or thousands of securities
- Stocks, bonds, options, swaps, structured products
- Equities, fixed income, foreign exchange, commodities

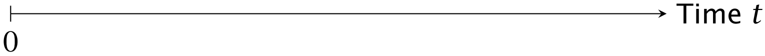
Security values at risk horizon τ

- Multiple underlying financial factors
- Financial model: distribution of factors at τ
- Security prices at τ in state ω
- Prices depend on cashflows from time τ to T
- Distribution of portfolio losses $L(\omega)$

Risk measure

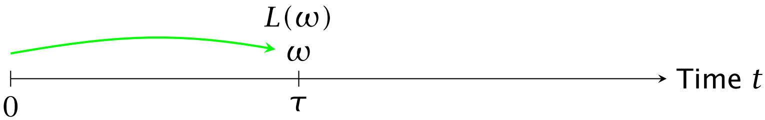
- Distribution of losses $L(\omega)$ is mapped to a risk measure $\rho(L)$

The Risk Measurement Problem



- Today: $t = 0$

The Risk Measurement Problem

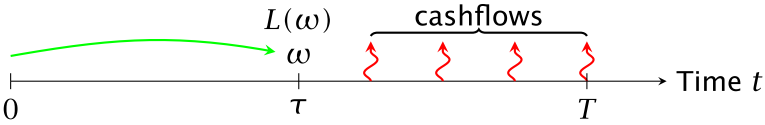


- Today: $t = 0$
- Risk horizon: $t = \tau$

$$\omega = \text{state at time } \tau$$

$$L(\omega) = \text{portfolio loss at time } \tau\text{, given state } \omega$$

The Risk Measurement Problem



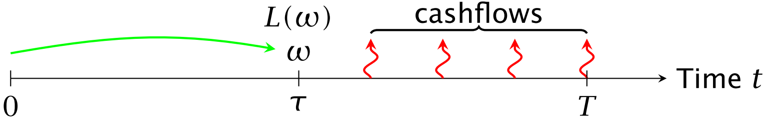
- Today: $t = 0$
- Risk horizon: $t = \tau$

$$\omega = \text{state at time } \tau$$

$$L(\omega) = \text{portfolio loss at time } \tau\text{, given state } \omega$$

- $L(\omega)$ depends on realized cashflows between τ and T

The Risk Measurement Problem



- Today: $t = 0$
- Risk horizon: $t = \tau$

$$\omega = \text{state at time } \tau$$

$$L(\omega) = \text{portfolio loss at time } \tau\text{, given state } \omega$$

- $L(\omega)$ depends on realized cashflows between τ and T
- Risk measure $\rho(L) \in \mathbb{R}$

Probability of large loss: $\mathsf{P}(L \geq c)$

$\mathrm{VAR}_\alpha(L) = \inf\{c \; : \; \mathsf{P}(L \geq c) \leq \alpha\}$

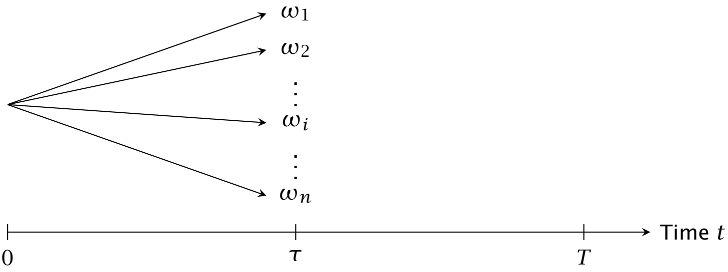
$\mathrm{CVAR}_\alpha(L) = \mathsf{E}[\,L \,|\, L \geq \mathrm{VAR}_\alpha(L)\,]$

Coherent risk measures ...

Related Literature

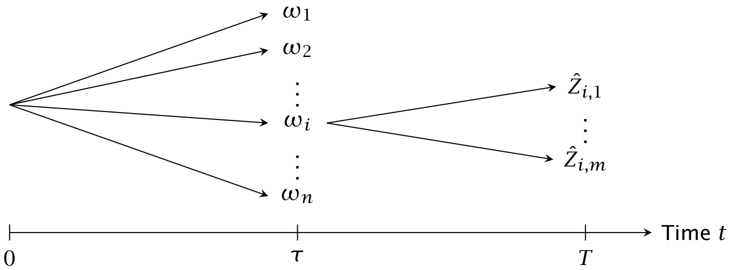
- Uniform nested simulation
  - Lee (1998)
  - Lee and Glynn (2003)
  - Gordy and Juneja (2006, 2008)
- Importance sampling
  - Glasserman, Heidelberger, Shahabuddin (2000)
- Stochastic kriging
  - Liu and Staum (2009)

The Risk Measurement Problem



- Simulate $\omega_1, \ldots, \omega_n$

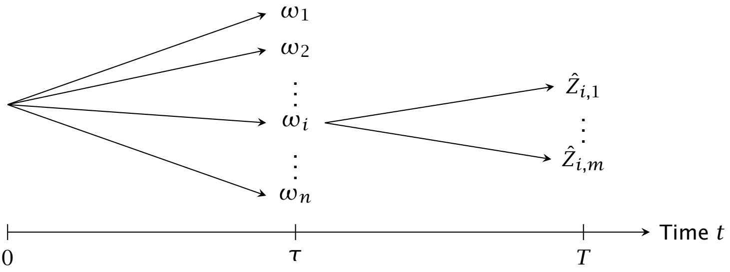
The Risk Measurement Problem



- Simulate $\omega_1, \ldots, \omega_n$
- For each ω_i: simulate future portfolio cashflows $\hat{Z}_{i,1}, \ldots, \hat{Z}_{i,m}$

$$\left.\hat{L}_i = \frac{1}{m} \sum_{j=1}^{m} \hat{Z}_{i,j} \right\} \text{ estimate of loss } L(\omega_i)$$

The Risk Measurement Problem



- Simulate $\omega_1, \ldots, \omega_n$
- For each ω_i: simulate future portfolio cashflows $\hat{Z}_{i,1}, \ldots, \hat{Z}_{i,m}$

$$\left. \hat{L}_i = \frac{1}{m} \sum_{j=1}^{m} \hat{Z}_{i,j} \right\} \text{ estimate of loss } L(\omega_i)$$

- Estimate probability of loss

$$\hat{\alpha} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{\hat{L}_i \geq c\}}$$

Probability of Loss: Gaussian Example

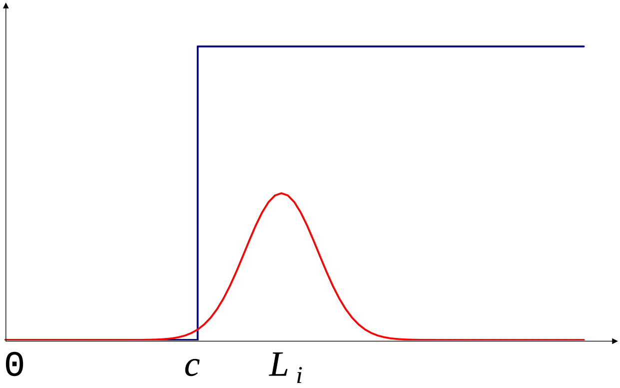
- First stage: $L(\omega_i) = \omega_i$, where $\omega_i \sim N(0, \sigma_1^2)$
- Second stage: $Z_{i,j} = \omega_i + \epsilon_{i,j}$, where $\epsilon_{i,j} \sim N(0, \sigma_2^2)$
- Probability of loss: $\alpha = \mathrm{P}(L \geq c) = \Phi(-c/\sigma_1)$

Estimator: $\hat{\alpha} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{\hat{L}_i \geq c\}}$ where $\hat{L}_i = L_i + \frac{1}{m}\sum_{j=1}^{m} \hat{Z}_{i,j}$

Mean-Squared Error (MSE):

$$\begin{aligned}\text{MSE} &= E[(\hat{\alpha} - \alpha)^2] \\ &= E[(\hat{\alpha} - E(\hat{\alpha}))^2] + (E[\hat{\alpha} - \alpha])^2 \\ &= \text{Variance} + \text{Bias}^2\end{aligned}$$

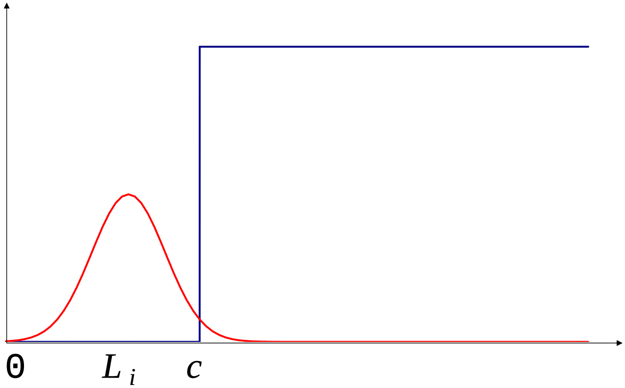
Bias Illustration



For $L_i > c$, $\mathbf{1}_{\{L_i \geq c\}} = 1$, but $E[\mathbf{1}_{\{\hat{L}_i \geq c\}}] = \mathrm{P}(\hat{L}_i \geq c) < 1$.

The local bias is negative: $E[\mathbf{1}_{\{\hat{L}_i \geq c\}} - 1] < 0$.

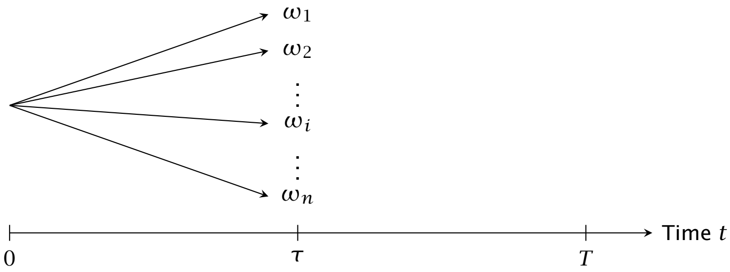
Bias Illustration



For $L_i < c$, $\mathbf{1}_{\{L_i \geq c\}} = 0$, but $E[\mathbf{1}_{\{\hat{L}_i \geq c\}}] = \mathrm{P}(\hat{L}_i \geq c) > 0$.

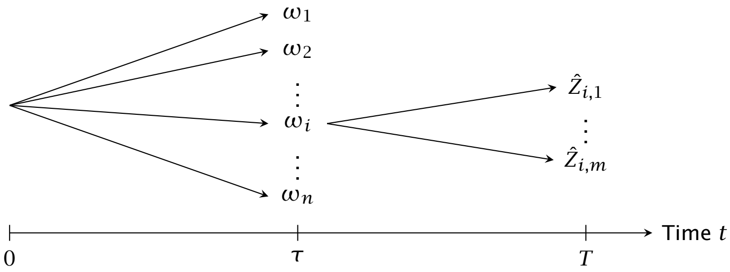
The local bias is positive: $E[\mathbf{1}_{\{\hat{L}_i \geq c\}} - 0] > 0$.

Optimal MSE Formulation



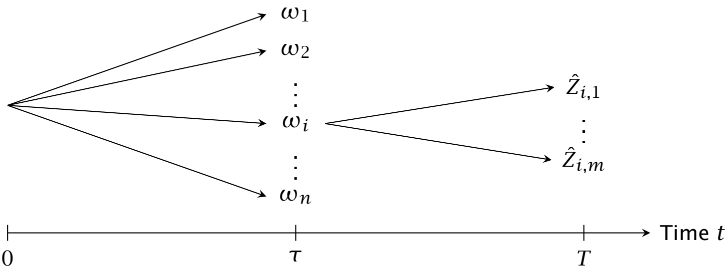
n first stage samples

Optimal MSE Formulation



n first stage samples $\qquad$ m second stage samples

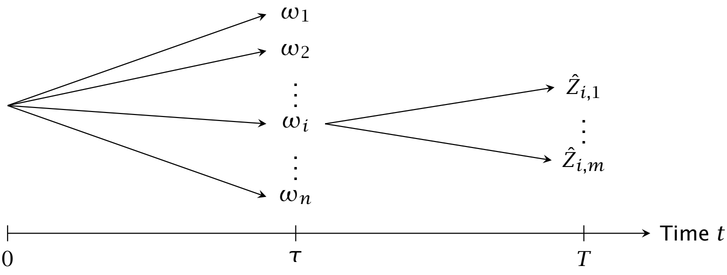
Optimal MSE Formulation



n first stage samples m second stage samples

total work: $k = mn$

Optimal MSE Formulation



n first stage samples m second stage samples

total work: $k = mn$

Optimal allocation problem:

$$\begin{array}{ll} \underset{n,m}{\text{minimize}} & \text{MSE} \\ \text{subject to} & nm = k \end{array}$$

Bias and Variance

$$\alpha = \mathsf{P}(L \geq c) \qquad \hat{\alpha} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{\hat{L}_i \geq c\}}$$

$$\mathsf{MSE} = \underbrace{\mathsf{E}\left[(\hat{\alpha} - \mathsf{E}\hat{\alpha})^2 \right]}_{\text{variance}} + \underbrace{\left(\mathsf{E}\left[\alpha - \hat{\alpha} \right] \right)^2}_{\text{bias}^2}$$

Bias and Variance

$$\alpha = \mathsf{P}(L \geq c) \qquad \hat{\alpha} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{\hat{L}_i \geq c\}}$$

$$\text{MSE} = \underbrace{\mathsf{E}\left[(\hat{\alpha} - \mathsf{E}\hat{\alpha})^2\right]}_{\text{variance}} + \underbrace{\left(\mathsf{E}\left[\alpha - \hat{\alpha}\right]\right)^2}_{\text{bias}^2}$$

Under mild technical assumptions, as $m, n \uparrow \infty$:

$$\text{variance} \to \frac{\alpha(1-\alpha)}{n} \qquad \text{bias} \to \frac{\gamma}{m}$$

Bias and Variance

$$\alpha = \mathsf{P}(L \geq c) \qquad \hat{\alpha} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{\hat{L}_i \geq c\}}$$

$$\text{MSE} = \underbrace{\mathsf{E}\left[(\hat{\alpha} - \mathsf{E}\hat{\alpha})^2\right]}_{\text{variance}} + \underbrace{\left(\mathsf{E}\left[\alpha - \hat{\alpha}\right]\right)^2}_{\text{bias}^2}$$

Under mild technical assumptions, as $m, n \uparrow \infty$:

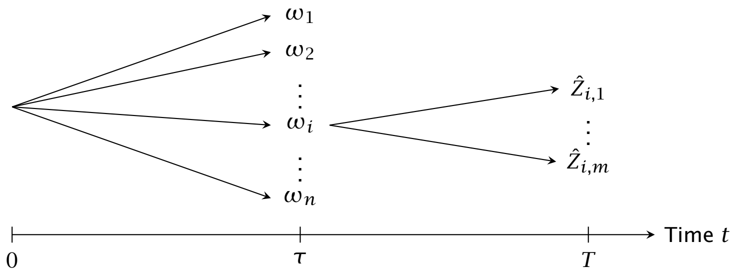
$$\text{variance} \to \frac{\alpha(1-\alpha)}{n} \qquad \text{bias} \to \frac{\gamma}{m}$$

Optimal allocation:

$$\begin{array}{ll} \underset{n,m}{\text{minimize}} & \text{MSE} \\ \text{subject to} & nm = k \end{array} \quad \Rightarrow \quad \begin{cases} n^* = Ck^{2/3} \\ m^* = \frac{1}{C}k^{1/3} \\ \text{MSE} \propto k^{-2/3} \end{cases}$$

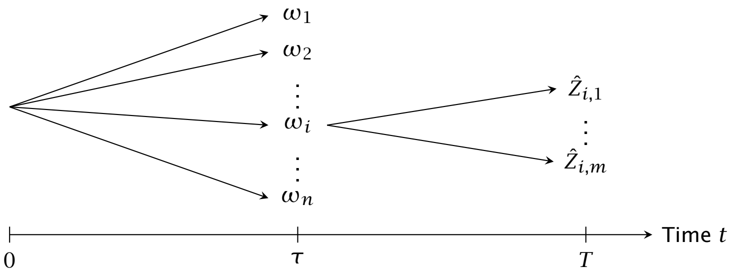
Gordy and Juneja (2006)

Optimal MSE Estimator



Optimal allocation: $n^* = Ck^{2/3}$, $m^* = \frac{1}{C}k^{1/3}$, MSE $\propto k^{-2/3}$

Optimal MSE Estimator

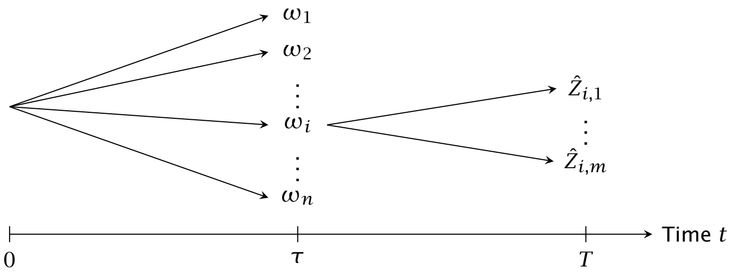


Optimal allocation: $n^* = Ck^{2/3}$, $m^* = \frac{1}{C}k^{1/3}$, MSE $\propto k^{-2/3}$

Observations:

- Similar expressions for VAR and CVAR, different constants

Optimal MSE Estimator

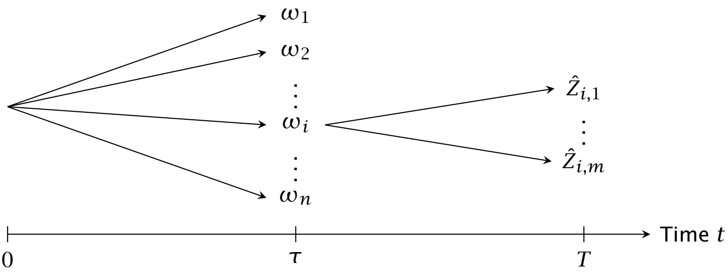


Optimal allocation: $n^* = Ck^{2/3}$, $m^* = \frac{1}{C}k^{1/3}$, MSE $\propto k^{-2/3}$

Observations:

- Similar expressions for VAR and CVAR, different constants
- Not clear how to implement! Need to estimate the constant C

Optimal MSE Estimator

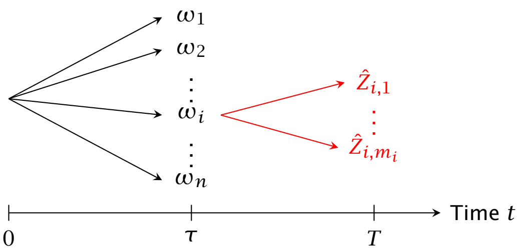


Optimal allocation: $n^* = Ck^{2/3}$, $m^* = \frac{1}{C}k^{1/3}$, MSE $\propto k^{-2/3}$

Observations:

- Similar expressions for VAR and CVAR, different constants
- Not clear how to implement! Need to estimate the constant C
- Can we do better?

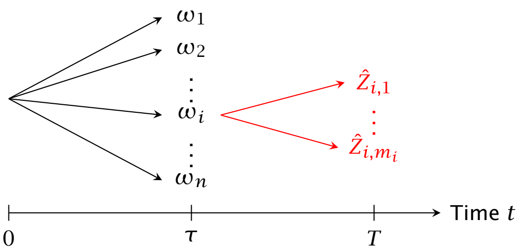
Non-Uniform Sampling



Idea: use a non-uniform number of stage 2 samples

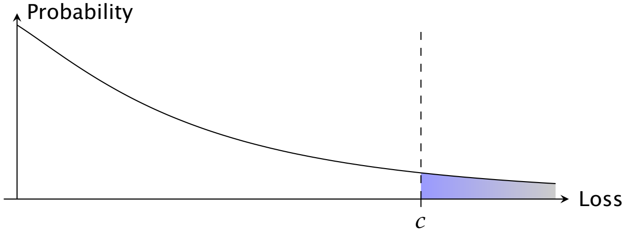
m_i = number of samples at ω_i

Non-Uniform Sampling



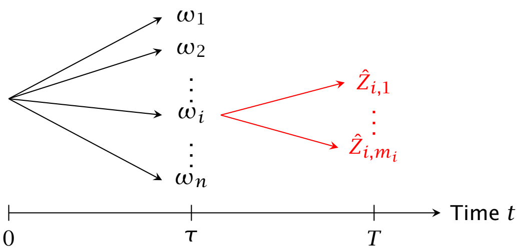
Idea: use a non-uniform number of stage 2 samples

m_i = number of samples at ω_i



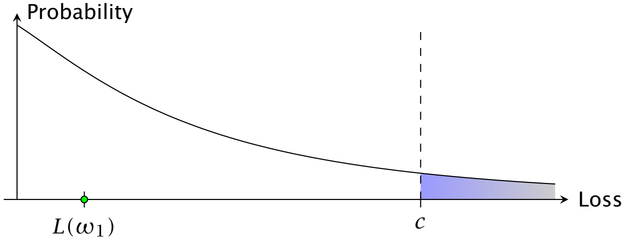
$$\hat{\alpha} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{\hat{L}_i \geq c\}}$$

Non-Uniform Sampling



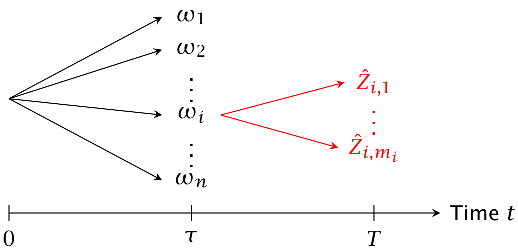
Idea: use a non-uniform number of stage 2 samples

m_i = number of samples at ω_i



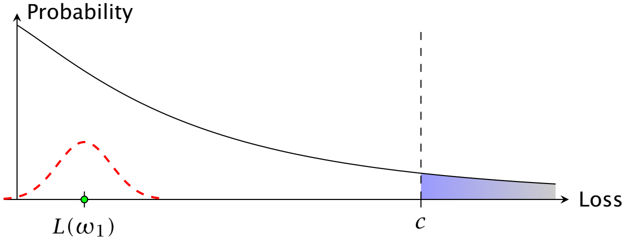
$$\hat{\alpha} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{\hat{L}_i \geq c\}}$$

Non-Uniform Sampling



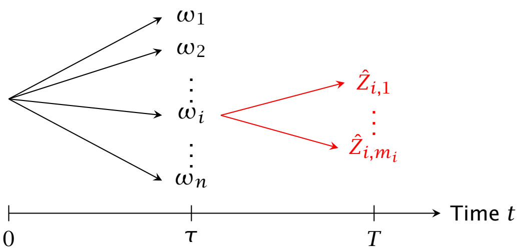
Idea: use a non-uniform number of stage 2 samples

m_i = number of samples at ω_i



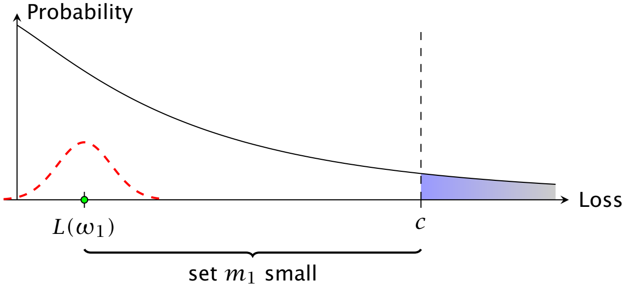
$$\hat{\alpha} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{\hat{L}_i \geq c\}}$$

Non-Uniform Sampling



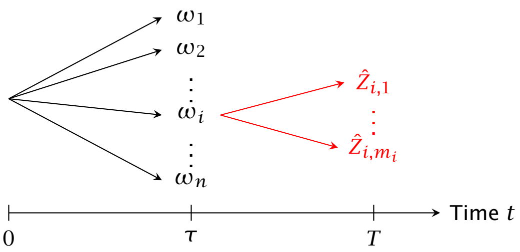
Idea: use a non-uniform number of stage 2 samples

m_i = number of samples at ω_i



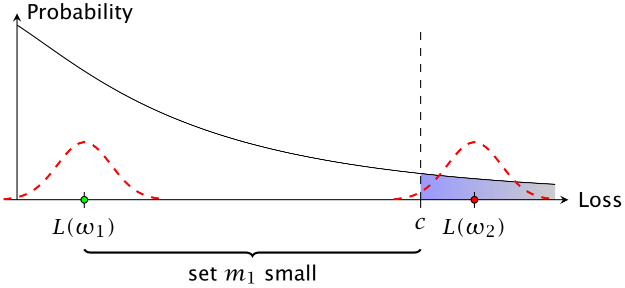
$$\hat{\alpha} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{\hat{L}_i \geq c\}}$$

Non-Uniform Sampling



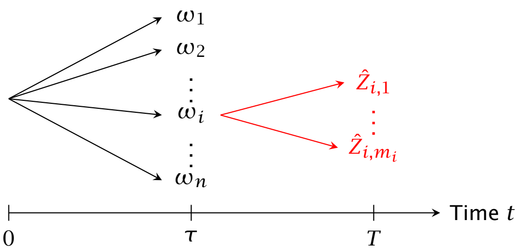
Idea: use a non-uniform number of stage 2 samples

m_i = number of samples at ω_i



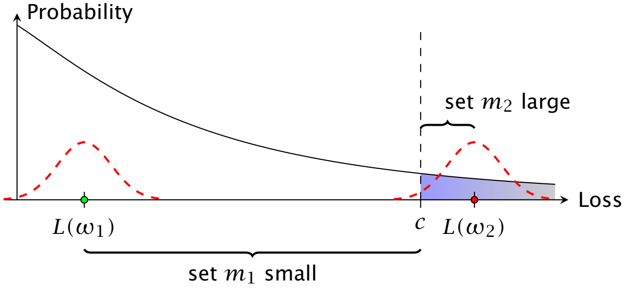
$$\hat{\alpha} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{\hat{L}_i \geq c\}}$$

Non-Uniform Sampling



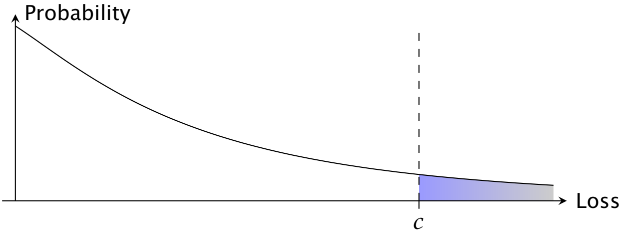
Idea: use a non-uniform number of stage 2 samples

m_i = number of samples at ω_i



$$\hat{\alpha} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{\hat{L}_i \geq c\}}$$

Stage 2 Algorithm



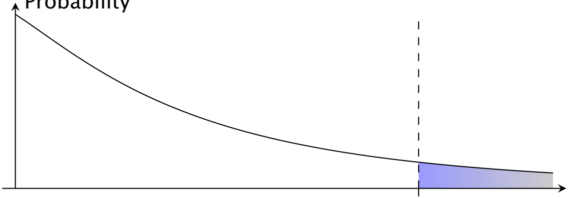
$$\hat{L}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} \hat{Z}_{i,j}$$

$$\hat{\alpha} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{\hat{L}_i \geq c\}}$$

Idea:

- Sequentially add stage 2 samples

Stage 2 Algorithm



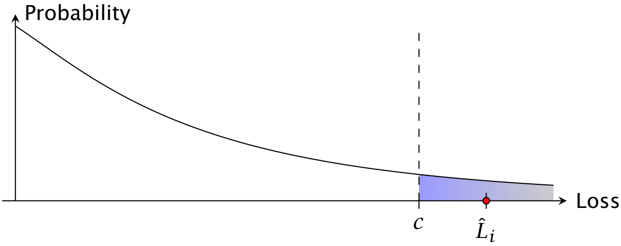
$$\hat{L}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} \hat{Z}_{i,j}$$

$$\hat{\alpha} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{\hat{L}_i \geq c\}}$$

Idea:

- Sequentially add stage 2 samples
- Add the next sample where it will most affect the estimate $\hat{\alpha}$

Stage 2 Algorithm



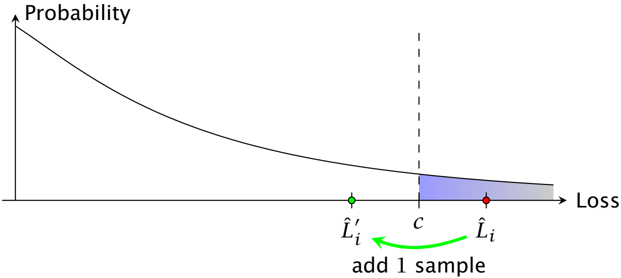
$$\hat{L}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} \hat{Z}_{i,j}$$

$$\hat{\alpha} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{\hat{L}_i \geq c\}}$$

Idea:

- Sequentially add stage 2 samples
- Add the next sample where it will most affect the estimate $\hat{\alpha}$

Stage 2 Algorithm



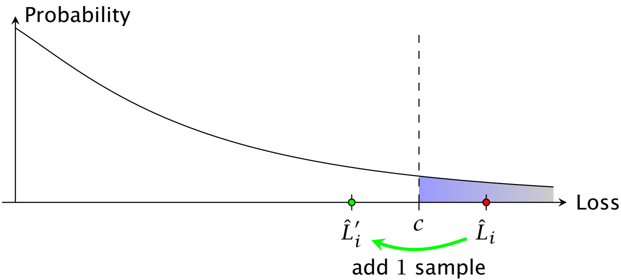
$$\hat{L}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} \hat{Z}_{i,j}$$

$$\hat{\alpha} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{\hat{L}_i \geq c\}}$$

Idea:

- Sequentially add stage 2 samples
- Add the next sample where it will most affect the estimate $\hat{\alpha}$

Stage 2 Algorithm



$$\hat{L}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} \hat{Z}_{i,j}$$

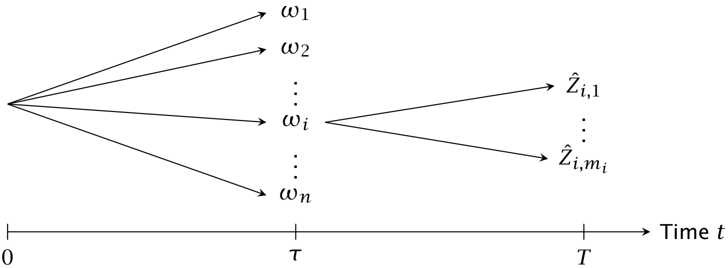
$$\hat{\alpha} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{\hat{L}_i \geq c\}}$$

Idea:

- Sequentially add stage 2 samples
- Add the next sample where it will most affect the estimate $\hat{\alpha}$
- Use a normal approximation: given one more sample at ω_i,

$$\mathrm{P}\,(\text{estimate } \hat{\alpha} \text{ changes}) \approx \Phi\left(-\frac{m_i}{\sigma_2}\left|\hat{L}_i - c\right|\right)$$

Non-Uniform Stage 2 Algorithm



- Simulate $\omega_1, \ldots, \omega_n$
- For each ℓ from 1 to k:

  Pick $i^* \in \operatorname*{argmin}_i \frac{m_i}{\sigma_2} \left| \hat{L}_i - c \right|$, Add 1 sample at ω_{i^*}

- Estimate probability of loss

$$\hat{\alpha} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{\hat{L}_i \geq c\}}$$

Key Result

Under suitable assumptions,

$$\text{bias} \propto \frac{1}{\bar{m}^2} \qquad \left(\text{vs. bias} \propto \frac{1}{m} \text{ under uniform sampling}\right)$$

Key Result

Under suitable assumptions,

$$\text{bias} \propto \frac{1}{\bar{m}^2} \qquad \left(\text{vs. bias} \propto \frac{1}{m} \text{ under uniform sampling}\right)$$

Proof Technique:
For a given ω_i, consider the sequential hypothesis testing problem:

- Observe IID samples $\hat{Z}_{i,1}, \hat{Z}_{i,2}, \ldots$ with $L(\omega_i) = \mathsf{E}[Z_{i,1}]$
- Hypotheses:

$$H_0(\omega_i) = \{L(\omega_i) < c\}$$

$$H_1(\omega_i) = \{L(\omega_i) \geq c\}$$

- We wish to determine which hypothesis is true, with a minimal number of observations

Our non-uniform sampling algorithm is solving many sequential hypothesis testing problems simultaneously

Rate of Convergence

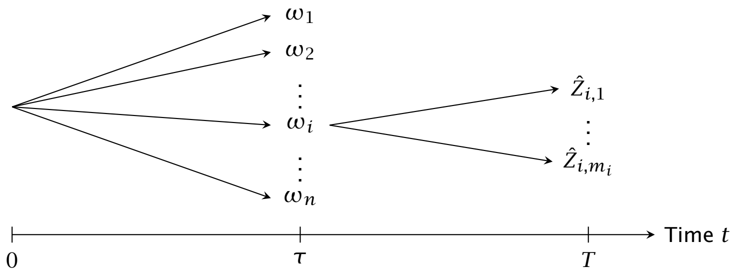
- Uniform algorithm:

$$\begin{array}{ll} \underset{n,m}{\text{minimize}} & \text{MSE} \\ \text{subject to} & nm = k \end{array} \quad \Rightarrow \quad \begin{cases} n^* \propto k^{2/3} \\ m^* \propto k^{1/3} \\ \text{MSE} \propto k^{-2/3} \end{cases}$$

- Non-uniform algorithm:

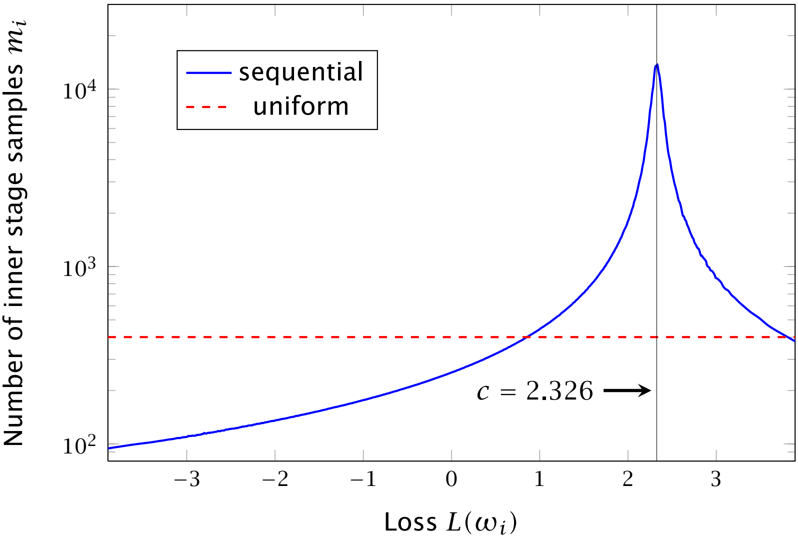
$$\begin{array}{ll} \underset{n,\bar{m}}{\text{minimize}} & \text{MSE} \\ \text{subject to} & n\bar{m} = k \end{array} \quad \Rightarrow \quad \begin{cases} n^* \propto k^{4/5} \\ \bar{m}^* \propto k^{1/5} \\ \text{MSE} \propto k^{-4/5} \end{cases}$$

Gaussian Example

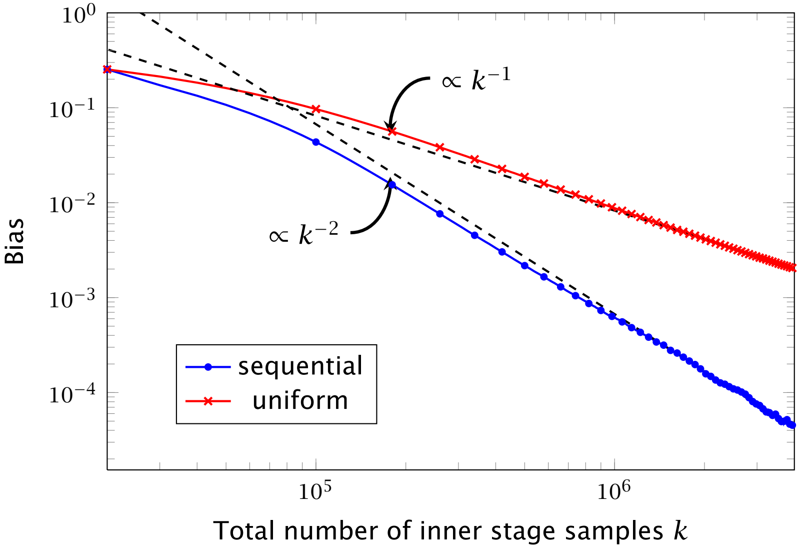


- First stage: $L(\omega_i) = \omega_i$, where $\omega_i \sim N(0, \sigma_1^2)$
- Second stage: $Z_{i,j} = \omega_i + \epsilon_{i,j}$, where $\epsilon_{i,j} \sim N(0, \sigma_2^2)$
- Probability of loss: $\mathsf{P}(L \geq c) = \Phi(-c/\sigma_1)$

Number of Inner Stage Samples versus Loss



Bias versus Number of Inner Stage Samples

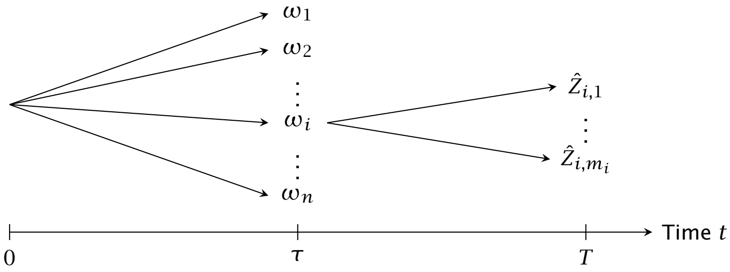


Numerical Results: Gaussian Example

$$\sigma_1 = 1,\ \sigma_2 = 5,\ \alpha = 0.1\%,\ k = 4{,}000{,}000$$

	n	$\bar{m}$	MSE	Rel MSE
$n = m = \sqrt{k}$	2,000	2,000	$5.7 \cdot 10^{-7}$	23
$n = k^{2/3},\ m = k^{1/3}$	25,200	159	$1.2 \cdot 10^{-6}$	48
uniform (optimal constant)	7,788	514	$2.5 \cdot 10^{-7}$	10
adaptive	30,628	132	$3.6 \cdot 10^{-8}$	1.5
optimal sequential	56,686	71	$2.5 \cdot 10^{-8}$	1

Put Option Example



- Stock price: $S_\tau(\omega) \triangleq S_0 e^{(\mu-\sigma^2/2)\tau+\sigma\sqrt{\tau}\omega}$
- $L(\omega) = X_0 - \mathrm{E}\left[e^{-r(T-\tau)} \max\left(K - S_T(\omega, W), 0\right) \,\middle|\, \omega \right]$ where

$$S_T(\omega, W) \triangleq S_\tau(\omega) e^{(r-\sigma^2/2)(T-\tau)+\sigma\sqrt{T-\tau}W}$$

and

$$\hat{Z}_{i,j} = X_0 - e^{-r(T-\tau)} \max\left(K - S_T(\omega_i, W_{i,j}), 0\right),$$

- Outer stage: the real-world distribution (μ)
- Inner stage: risk-neutral distribution (r)

Numerical Results: Put Option

$$S_0 = 100,\ K = 95,\ \sigma = 20\%,\ \tau = 1/52,\ T = 0.25$$

$$\alpha = 0.1\%,\ k = 4{,}000{,}000$$

	n	$\bar{m}$	MSE	Rel MSE
$n = m = \sqrt{k}$	2,000	2,000	$5.6 \cdot 10^{-7}$	12
$n = k^{2/3},\ m = k^{1/3}$	25,200	159	$8.2 \cdot 10^{-6}$	175
uniform (optimal constant)	2,570	1,556	$4.8 \cdot 10^{-7}$	10
adaptive	14,384	284	$9.2 \cdot 10^{-8}$	2
optimal sequential	26,508	151	$4.7 \cdot 10^{-8}$	1

Summary

- Nested simulation can provide a more realistic assessment of risk

- Reduced computational burden by
  - Non-uniform inner sampling to reduce bias
  - More outer sampling to reduce variance

- MSE reduced by factors from 4 to over 100