Monte Carlo methods in credit risk management

Monte Carlo methods in credit risk management

Let P be a credit portfolio consisting of m credits.
The loss function is $L = \sum_{i=1}^{m} L_i$ and the single credit losses L_i are independent conditioned on a vector Z of economical impact factors.

Monte Carlo methods in credit risk management

Let P be a credit portfolio consisting of m credits.
The loss function is $L = \sum_{i=1}^{m} L_i$ and the single credit losses L_i are independent conditioned on a vector Z of economical impact factors.

Goal: Determine $\mathrm{VaR}_\alpha(L) = q_\alpha(L)$, $\mathrm{CVaR}_\alpha = E(L|L > q_\alpha(L))$, $\mathrm{CVaR}_{i,\alpha} = E(L_i|L > q_\alpha(L))$, for all i.

Monte Carlo methods in credit risk management

Let P be a credit portfolio consisting of m credits.
The loss function is $\mathrm{L} = \sum_{\mathrm{i}=1}^{\mathrm{m}} \mathrm{L_i}$ and the single credit losses $\mathrm{L_i}$ are independent conditioned on a vector Z of economical impact factors.

Goal: Determine $\mathrm{VaR}_\alpha(\mathrm{L}) = \mathrm{q}_\alpha(\mathrm{L})$, $\mathrm{CVaR}_\alpha = \mathrm{E}(\mathrm{L}|\mathrm{L} > \mathrm{q}_\alpha(\mathrm{L}))$, $\mathrm{CVaR}_{\mathrm{i},\alpha} = \mathrm{E}(\mathrm{L_i}|\mathrm{L} > \mathrm{q}_\alpha(\mathrm{L}))$, for all i.

Application of Monte Carlo (MC) simulation has to deal with the simulation of rare events!
E.g. for $\alpha = 0,99$ only 1% of the standard MC simulations will lead to a loss L, such that $\mathrm{L} > \mathrm{q}_\alpha(\mathrm{L})$.

Monte Carlo methods in credit risk management

Let P be a credit portfolio consisting of m credits.
The loss function is $L = \sum_{i=1}^{m} L_i$ and the single credit losses L_i are independent conditioned on a vector Z of economical impact factors.

Goal: Determine $VaR_\alpha(L) = q_\alpha(L)$, $CVaR_\alpha = E(L|L > q_\alpha(L))$, $CVaR_{i,\alpha} = E(L_i|L > q_\alpha(L))$, for all i.

Application of Monte Carlo (MC) simulation has to deal with the simulation of rare events!
E.g. for $\alpha = 0,99$ only 1% of the standard MC simulations will lead to a loss L, such that $L > q_\alpha(L)$.

The standard MC estimator is:

$$\widehat{CVaR}_\alpha^{(MC)}(L) = \frac{1}{\sum_{i=1}^{n} I_{(q_\alpha,+\infty)}(L^{(i)})} \sum_{i=1}^{n} L^{(i)} I_{(q_\alpha,+\infty)}(L^{(i)}) ,$$

where L_i is the value of the loss in the ith simulation run.

Monte Carlo methods in credit risk management

Let P be a credit portfolio consisting of m credits.
The loss function is $L = \sum_{i=1}^{m} L_i$ and the single credit losses L_i are independent conditioned on a vector Z of economical impact factors.

Goal: Determine $\mathrm{VaR}_\alpha(L) = q_\alpha(L)$, $\mathrm{CVaR}_\alpha = E(L|L > q_\alpha(L))$, $\mathrm{CVaR}_{i,\alpha} = E(L_i|L > q_\alpha(L))$, for all i.

Application of Monte Carlo (MC) simulation has to deal with the simulation of rare events!
E.g. for $\alpha = 0,99$ only 1% of the standard MC simulations will lead to a loss L, such that $L > q_\alpha(L)$.

The standard MC estimator is:

$$\widehat{\mathrm{CVaR}}_\alpha^{(MC)}(L) = \frac{1}{\sum_{i=1}^{n} I_{(q_\alpha,+\infty)}(L^{(i)})} \sum_{i=1}^{n} L^{(i)} I_{(q_\alpha,+\infty)}(L^{(i)}) \, ,$$

where L_i is the value of the loss in the ith simulation run.
$\widehat{\mathrm{CVaR}}_\alpha^{(MC)}(L)$ is unstable, i.e. it has a very high variance, if the number of simulation runs ist not very high.

Basics of importance sampling

Basics of importance sampling

Let X be a r.v. in a probability space $(\Omega, \mathcal{F}, \mathrm{P})$ with absolutely continuous distribution function and density function f.

Goal: Determine $\theta = \mathrm{E}(\mathrm{h}(\mathrm{X})) = \int_{-\infty}^{\infty} \mathrm{h}(\mathrm{x})\mathrm{f}(\mathrm{x})\mathrm{dx}$ for some given function h.

Basics of importance sampling

Let X be a r.v. in a probability space $(\Omega, \mathcal{F}, P)$ with absolutely continuous distribution function and density function f.

Goal: Determine $\theta = E(h(X)) = \int_{-\infty}^{\infty} h(x)f(x)dx$ for some given function h.

Examples:
Set $h(x) = I_A(x)$ to compute the probability of an event A.

Set $h(x) = xI_{X>c}(x)$ with $c = VaR(X)$ to compute $CVaR(X)$.

Basics of importance sampling

Let X be a r.v. in a probability space $(\Omega, \mathcal{F}, P)$ with absolutely continuous distribution function and density function f.

Goal: Determine $\theta = E(h(X)) = \int_{-\infty}^{\infty} h(x)f(x)dx$ for some given function h.

Examples:
Set $h(x) = I_A(x)$ to compute the probability of an event A.

Set $h(x) = xI_{X>c}(x)$ with $c = VaR(X)$ to compute $CVaR(X)$.

Algorithm: Monte Carlo integration

(1) Simulate $X_1, X_2, \ldots, X_n$ independently with density f.

(2) Compute the standard MC estimator $\hat{\theta}_n^{(MC)} = \frac{1}{n}\sum_{i=1}^{n} h(X_i)$.

Basics of importance sampling

Let X be a r.v. in a probability space $(\Omega, \mathcal{F}, P)$ with absolutely continuous distribution function and density function f.

Goal: Determine $\theta = E(h(X)) = \int_{-\infty}^{\infty} h(x)f(x)dx$ for some given function h.

Examples:
Set $h(x) = I_A(x)$ to compute the probability of an event A.

Set $h(x) = xI_{x>c}(x)$ with $c = VaR(X)$ to compute $CVaR(X)$.

Algorithm: Monte Carlo integration

(1) Simulate $X_1, X_2, \ldots, X_n$ independently with density f.

(2) Compute the standard MC estimator $\hat{\theta}_n^{(MC)} = \frac{1}{n}\sum_{i=1}^{n} h(X_i)$.

The strong low of large numbers implies $\lim_{n\to\infty} \hat{\theta}_n^{(MC)} = \theta$ almost surely.

Basics of importance sampling

Let X be a r.v. in a probability space $(\Omega, \mathcal{F}, P)$ with absolutely continuous distribution function and density function f.

Goal: Determine $\theta = E(h(X)) = \int_{-\infty}^{\infty} h(x)f(x)dx$ for some given function h.

Examples:
Set $h(x) = I_A(x)$ to compute the probability of an event A.

Set $h(x) = xI_{x>c}(x)$ with $c = VaR(X)$ to compute $CVaR(X)$.

Algorithm: Monte Carlo integration

(1) Simulate X_1,X_2,. . . , X_n independently with density f.

(2) Compute the standard MC estimator $\hat{\theta}_n^{(MC)} = \frac{1}{n}\sum_{i=1}^{n} h(X_i)$.

The strong low of large numbers implies $\lim_{n\to\infty} \hat{\theta}_n^{(MC)} = \theta$ almost surely.

In case of rare events, e.g. $h(x) = I_A(x)$ with $P(A) << 1$, the convergence is very slow.

Importance sampling (contd.)

Importance sampling (contd.)

Let g be a probability density function, such that $f(x) > 0 \Rightarrow g(x) > 0$.

We define the *likelihood ratio* as: $r(x) := \begin{cases} \frac{f(x)}{g(x)} & g(x) > 0 \\ 0 & g(x) = 0 \end{cases}$

Importance sampling (contd.)

Let g be a probability density function, such that $f(x) > 0 \Rightarrow g(x) > 0$.

We define the *likelihood ratio* as: $r(x) := \begin{cases} \frac{f(x)}{g(x)} & g(x) > 0 \\ 0 & g(x) = 0 \end{cases}$

The following equality holds:

$$\theta = \int_{-\infty}^{\infty} h(x)r(x)g(x)dx = E_g(h(x)r(x))$$

Algorithm: Importance sampling

(1) Simulate $X_1, X_2, \ldots, X_n$ independently with density g.

(2) Compute the IS-estimator $\hat{\theta}_n^{(IS)} = \frac{1}{n}\sum_{i=1}^{n} h(X_i)r(X_i)$.

g is called *importance sampling density* (IS density).

Importance sampling (contd.)

Let g be a probability density function, such that $f(x) > 0 \Rightarrow g(x) > 0$.

We define the *likelihood ratio* as: $r(x) := \begin{cases} \frac{f(x)}{g(x)} & g(x) > 0 \\ 0 & g(x) = 0 \end{cases}$

The following equality holds:

$$\theta = \int_{-\infty}^{\infty} h(x)r(x)g(x)dx = E_g(h(x)r(x))$$

Algorithm: Importance sampling

(1) Simulate $X_1, X_2, \ldots, X_n$ independently with density g.

(2) Compute the IS-estimator $\hat{\theta}_n^{(IS)} = \frac{1}{n}\sum_{i=1}^{n} h(X_i)r(X_i)$.

g is called *importance sampling density* (IS density).

Goal: choose an IS density g such that the variance of the IS estimator is much smaller than the variance of the standard MC-estimator.

$$\operatorname{var}\left(\hat{\theta}_n^{(IS)}\right) = \frac{1}{n^2}(E_g(h^2(X)r^2(X)) - \theta^2)$$

$$\operatorname{var}\left(\hat{\theta}_n^{(MC)}\right) = \frac{1}{n^2}(E(h^2(X)) - \theta^2)$$

Importance sampling (contd.)

Importance sampling (contd.)

Theoretically the variance of the IS estimator can be reduced to 0!

Importance sampling (contd.)

Theoretically the variance of the IS estimator can be reduced to 0!

Assume $\mathrm{h}(\mathrm{x}) \geq 0, \forall \mathrm{x}$.

For $\mathrm{g}^*(\mathrm{x}) = \mathrm{f}(\mathrm{x})\mathrm{h}(\mathrm{x})/\mathrm{E}(\mathrm{h}(\mathrm{x}))$ we get : $\hat{\theta}_1^{(\mathrm{IS})} = \mathrm{h}(\mathrm{X}_1)\mathrm{r}(\mathrm{X}_1) = \mathrm{E}(\mathrm{h}(\mathrm{X}))$.

The IS estimator yields the correct value already after a single simulation!

Importance sampling (contd.)

Theoretically the variance of the IS estimator can be reduced to 0!

Assume $h(x) \geq 0, \forall x$.

For $g^*(x) = f(x)h(x)/E(h(x))$ we get : $\hat{\theta}_1^{(IS)} = h(X_1)r(X_1) = E(h(X))$.

The IS estimator yields the correct value already after a single simulation!

Let $h(x) = I_{\{X \geq c\}}(x)$ where $c >> E(X)$ (rare event).

Importance sampling (contd.)

Theoretically the variance of the IS estimator can be reduced to 0!

Assume $h(x) \geq 0, \forall x$.

For $g^*(x) = f(x)h(x)/E(h(x))$ we get : $\hat{\theta}_1^{(IS)} = h(X_1)r(X_1) = E(h(X))$.

The IS estimator yields the correct value already after a single simulation!

Let $h(x) = I_{\{X \geq c\}}(x)$ where $c >> E(X)$ (rare event).

We have $E(h^2(X)) = P(X \geq c)$ and

$$E_g(h^2(X)r^2(X)) = \int_{-\infty}^{\infty} h^2(x)r^2(x)g(x)dx = E_g(r^2(X); X \geq c) =$$

$$\int_{-\infty}^{\infty} h^2(x)r(x)f(x)dx = \int_{-\infty}^{\infty} h(x)r(x)f(x)dx = E_f(r(X); X \geq c)$$

Importance sampling (contd.)

Theoretically the variance of the IS estimator can be reduced to 0!

Assume $h(x) \geq 0, \forall x$.

For $g^*(x) = f(x)h(x)/E(h(x))$ we get : $\hat{\theta}_1^{(IS)} = h(X_1)r(X_1) = E(h(X))$.

The IS estimator yields the correct value already after a single simulation!

Let $h(x) = I_{\{X \geq c\}}(x)$ where $c >> E(X)$ (rare event).

We have $E(h^2(X)) = P(X \geq c)$ and

$$E_g(h^2(X)r^2(X)) = \int_{-\infty}^{\infty} h^2(x)r^2(x)g(x)dx = E_g(r^2(X); X \geq c) =$$

$$\int_{-\infty}^{\infty} h^2(x)r(x)f(x)dx = \int_{-\infty}^{\infty} h(x)r(x)f(x)dx = E_f(r(X); X \geq c)$$

Goal: choose g such that $E_g(h^2(X)r^2(X))$ becomes small, i.e. such that $r(x)$ is small for $x \geq c$.

Importance sampling (contd.)

Theoretically the variance of the IS estimator can be reduced to 0!

Assume $h(x) \geq 0, \forall x$.

For $g^*(x) = f(x)h(x)/E(h(x))$ we get : $\hat{\theta}_1^{(IS)} = h(X_1)r(X_1) = E(h(X))$.

The IS estimator yields the correct value already after a single simulation!

Let $h(x) = I_{\{X \geq c\}}(x)$ where $c >> E(X)$ (rare event).

We have $E(h^2(X)) = P(X \geq c)$ and

$$E_g(h^2(X)r^2(X)) = \int_{-\infty}^{\infty} h^2(x)r^2(x)g(x)dx = E_g(r^2(X); X \geq c) =$$

$$\int_{-\infty}^{\infty} h^2(x)r(x)f(x)dx = \int_{-\infty}^{\infty} h(x)r(x)f(x)dx = E_f(r(X); X \geq c)$$

Goal: choose g such that $E_g(h^2(X)r^2(X))$ becomes small, i.e. such that $r(x)$ is small for $x \geq c$. Aquivalently, the event $X \geq c$ should be more probable under density g than under density f.

Exponential tilting: Determining the IS density for light tailed r.v.

Exponential tilting: Determining the IS density for light tailed r.v.

Let $M_X(t) : \mathbb{R} \to \mathbb{R}$ be the moment generating function of the r.v. X with probability density f:

$$M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x)dx$$

Exponential tilting: Determining the IS density for light tailed r.v.

Let $M_X(t)\colon \mathbb{R} \to \mathbb{R}$ be the moment generating function of the r.v. X with probability density f:

$$M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x) dx$$

Consider the IS density $g_t(x) := \frac{e^{tx} f(x)}{M_X(t)}$. Then $r_t(x) = \frac{f(x)}{g_t(x)} = M_X(t) e^{-tx}$.

Exponential tilting: Determining the IS density for light tailed r.v.

Let $M_X(t) \colon \mathbb{R} \to \mathbb{R}$ be the moment generating function of the r.v. X with probability density f:

$$M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x) dx$$

Consider the IS density $g_t(x) := \frac{e^{tx} f(x)}{M_X(t)}$. Then
$r_t(x) = \frac{f(x)}{g_t(x)} = M_X(t) e^{-tx}$.
Let $\mu_t := E_{g_t}(X) = E(Xe^{tX})/M_X(t)$.

Exponential tilting: Determining the IS density for light tailed r.v.

Let $M_X(t)\colon \mathbb{R} \to \mathbb{R}$ be the moment generating function of the r.v. X with probability density f:

$$M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x)dx$$

Consider the IS density $g_t(x) := \frac{e^{tx} f(x)}{M_X(t)}$. Then $r_t(x) = \frac{f(x)}{g_t(x)} = M_X(t)e^{-tx}$.

Let $\mu_t := E_{g_t}(X) = E(Xe^{tX})/M_X(t)$.

How to determine a suitable t for a specific $h(x)$?

For example for the estimation of the tail probability?

Exponential tilting: Determining the IS density for light tailed r.v.

Let $M_X(t)\colon \mathbb{R} \to \mathbb{R}$ be the moment generating function of the r.v. X with probability density f:

$$M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x) dx$$

Consider the IS density $g_t(x) := \frac{e^{tx} f(x)}{M_X(t)}$. Then
$r_t(x) = \frac{f(x)}{g_t(x)} = M_X(t) e^{-tx}$.
Let $\mu_t := E_{g_t}(X) = E(Xe^{tX})/M_X(t)$.
How to determine a suitable t for a specific $h(x)$?
For example for the estimation of the tail probability?
Goal: choose t such that $E(r(X); X \geq c) = E(I_{X \geq c} M_X(t) e^{-tX})$ becomes small.

Exponential tilting: Determining the IS density for light tailed r.v.

Let $M_X(t)\colon \mathbb{R} \to \mathbb{R}$ be the moment generating function of the r.v. X with probability density f:

$$M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x)dx$$

Consider the IS density $g_t(x) := \frac{e^{tx} f(x)}{M_X(t)}$. Then
$r_t(x) = \frac{f(x)}{g_t(x)} = M_X(t)e^{-tx}$.
Let $\mu_t := E_{g_t}(X) = E(Xe^{tX})/M_X(t)$.
How to determine a suitable t for a specific $h(x)$?
For example for the estimation of the tail probability?
Goal: choose t such that $E(r(X); X \geq c) = E(I_{X\geq c} M_X(t)e^{-tX})$ becomes small.
$e^{-tx} \leq e^{-tc}$, for $x \geq c$, $t \geq 0 \Rightarrow E(I_{X\geq c} M_X(t)e^{-tX}) \leq M_X(t)e^{-tc}$.

Exponential tilting: Determining the IS density for light tailed r.v.

Let $M_X(t) \colon \mathbb{R} \to \mathbb{R}$ be the moment generating function of the r.v. X with probability density f:

$$M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x) dx$$

Consider the IS density $g_t(x) := \frac{e^{tx} f(x)}{M_X(t)}$. Then
$r_t(x) = \frac{f(x)}{g_t(x)} = M_X(t) e^{-tx}$.
Let $\mu_t := E_{g_t}(X) = E(Xe^{tX})/M_X(t)$.
How to determine a suitable t for a specific $h(x)$?
For example for the estimation of the tail probability?
Goal: choose t such that $E(r(X); X \geq c) = E(I_{X \geq c} M_X(t) e^{-tX})$ becomes small.
$e^{-tx} \leq e^{-tc}$, for $x \geq c$, $t \geq 0 \Rightarrow E(I_{X \geq c} M_X(t) e^{-tX}) \leq M_X(t) e^{-tc}$.
Set $t = \operatorname{argmin}\{M_X(t) e^{-tc} \colon t \geq 0\}$ which imples $t = t(c)$, where $t(c)$ is the solution of the equation $\mu_t = c$.

Exponential tilting: Determining the IS density for light tailed r.v.

Let $M_X(t) : \mathbb{R} \to \mathbb{R}$ be the moment generating function of the r.v. X with probability density f:

$$M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x) dx$$

Consider the IS density $g_t(x) := \frac{e^{tx} f(x)}{M_X(t)}$. Then
$r_t(x) = \frac{f(x)}{g_t(x)} = M_X(t) e^{-tx}$.
Let $\mu_t := E_{g_t}(X) = E(Xe^{tX})/M_X(t)$.
How to determine a suitable t for a specific $h(x)$?
For example for the estimation of the tail probability?
Goal: choose t such that $E(r(X); X \geq c) = E(I_{X \geq c} M_X(t) e^{-tX})$ becomes small.
$e^{-tx} \leq e^{-tc}$, for $x \geq c$, $t \geq 0 \Rightarrow E(I_{X \geq c} M_X(t) e^{-tX}) \leq M_X(t) e^{-tc}$.
Set $t = \operatorname{argmin}\{M_X(t) e^{-tc} : t \geq 0\}$ which imples $t = t(c)$, where $t(c)$ is the solution of the equation $\mu_t = c$.

(A unique solution of the above equality exists for all relevant values of c, see e.g. Embrechts et al. for a proof).

IS in the case of probability measures

(useful for the estimation of the credit portfolio risk)

IS in the case of probability measures

(useful for the estimation of the credit portfolio risk)

Let f and g be probability densities. Define probability measures P and Q:
$P(A) := \int_{x \in A} f(x) dx$ and $Q(A) := \int_{x \in A} g(x) dx$ for $A \subset \mathbb{R}$.

IS in the case of probability measures

(useful for the estimation of the credit portfolio risk)

Let f and g be probability densities. Define probability measures P and Q:
$\mathrm{P}(\mathrm{A}) := \int_{\mathrm{x}\in\mathrm{A}} \mathrm{f}(\mathrm{x})\mathrm{dx}$ and $\mathrm{Q}(\mathrm{A}) := \int_{\mathrm{x}\in\mathrm{A}} \mathrm{g}(\mathrm{x})\mathrm{dx}$ for $\mathrm{A} \subset \mathbb{R}$.

Goal: Estimate the expected value $\theta := \mathrm{E}^{\mathrm{P}}(\mathrm{h}(\mathrm{X}))$ of a given function $\mathrm{h}: \mathcal{F} \to \mathbb{R}$ in the probability space $(\Omega, \mathcal{F}, \mathrm{P})$.

IS in the case of probability measures

(useful for the estimation of the credit portfolio risk)

Let f and g be probability densities. Define probability measures P and Q:
$P(A) := \int_{x \in A} f(x)dx$ and $Q(A) := \int_{x \in A} g(x)dx$ for $A \subset \mathbb{R}$.

Goal: Estimate the expected value $\theta := E^{P}(h(X))$ of a given function $h: \mathcal{F} \to \mathbb{R}$ in the probability space $(\Omega, \mathcal{F}, P)$.

We have $\theta := E^{P}(h(X)) = E^{Q}(h(X)r(X))$ with $r(x) := dP/dQ$, thus r is the density of P w.r.t. Q.

IS in the case of probability measures

(useful for the estimation of the credit portfolio risk)

Let f and g be probability densities. Define probability measures P and Q:

$P(A) := \int_{x \in A} f(x)dx$ and $Q(A) := \int_{x \in A} g(x)dx$ for $A \subset \mathbb{R}$.

Goal: Estimate the expected value $\theta := E^P(h(X))$ of a given function $h: \mathcal{F} \to \mathbb{R}$ in the probability space $(\Omega, \mathcal{F}, P)$.

We have $\theta := E^P(h(X)) = E^Q(h(X)r(X))$ with $r(x) := dP/dQ$, thus r is the density of P w.r.t. Q.

Exponential tilting in the case of probability measures:

Let X be a r.v. in $(\Omega, \mathcal{F}, P)$ such that $M_X(t) = E^P(\exp\{tX\}) < \infty$, $\forall t$.

IS in the case of probability measures

(useful for the estimation of the credit portfolio risk)

Let f and g be probability densities. Define probability measures P and Q:
$P(A) := \int_{x \in A} f(x)dx$ and $Q(A) := \int_{x \in A} g(x)dx$ for $A \subset \mathbb{R}$.
Goal: Estimate the expected value $\theta := E^P(h(X))$ of a given function $h: \mathcal{F} \to \mathbb{R}$ in the probability space $(\Omega, \mathcal{F}, P)$.
We have $\theta := E^P(h(X)) = E^Q(h(X)r(X))$ with $r(x) := dP/dQ$, thus r is the density of P w.r.t. Q.

Exponential tilting in the case of probability measures:
Let X be a r.v. in $(\Omega, \mathcal{F}, P)$ such that $M_X(t) = E^P(\exp\{tX\}) < \infty$, $\forall t$.
Define a probability measure Q_t in $(\Omega, \mathcal{F})$, such that
$dQ_t/dP = \exp(tX)/M_X(t)$, i.e. $Q_t(A) := E^P\left(\frac{\exp\{tX\}}{M_X(t)}; A\right)$.

IS in the case of probability measures

(useful for the estimation of the credit portfolio risk)

Let f and g be probability densities. Define probability measures P and Q:

$P(A) := \int_{x\in A} f(x)dx$ and $Q(A) := \int_{x\in A} g(x)dx$ for $A \subset \mathbb{R}$.

Goal: Estimate the expected value $\theta := E^P(h(X))$ of a given function $h\colon \mathcal{F} \to \mathbb{R}$ in the probability space $(\Omega, \mathcal{F}, P)$.

We have $\theta := E^P(h(X)) = E^Q(h(X)r(X))$ with $r(x) := dP/dQ$, thus r is the density of P w.r.t. Q.

Exponential tilting in the case of probability measures:

Let X be a r.v. in $(\Omega, \mathcal{F}, P)$ such that $M_X(t) = E^P(\exp\{tX\}) < \infty$, $\forall t$.

Define a probability measure Q_t in $(\Omega, \mathcal{F})$, such that

$dQ_t/dP = \exp(tX)/M_X(t)$, i.e. $Q_t(A) := E^P\left(\frac{\exp\{tX\}}{M_X(t)}; A\right)$.

We have $\frac{dP}{dQ_t} = M_X(t)\exp(-tX) =: r_t(X)$.

IS in the case of probability measures

(useful for the estimation of the credit portfolio risk)

Let f and g be probability densities. Define probability measures P and Q:

$P(A) := \int_{x \in A} f(x)dx$ and $Q(A) := \int_{x \in A} g(x)dx$ for $A \subset \mathbb{R}$.

Goal: Estimate the expected value $\theta := E^P(h(X))$ of a given function $h \colon \mathcal{F} \to \mathbb{R}$ in the probability space $(\Omega, \mathcal{F}, P)$.

We have $\theta := E^P(h(X)) = E^Q(h(X)r(X))$ with $r(x) := dP/dQ$, thus r is the density of P w.r.t. Q.

Exponential tilting in the case of probability measures:

Let X be a r.v. in $(\Omega, \mathcal{F}, P)$ such that $M_X(t) = E^P(\exp\{tX\}) < \infty$, $\forall t$.

Define a probability measure Q_t in $(\Omega, \mathcal{F})$, such that $dQ_t/dP = \exp(tX)/M_X(t)$, i.e. $Q_t(A) := E^P\left(\frac{\exp\{tX\}}{M_X(t)}; A\right)$.

We have $\frac{dP}{dQ_t} = M_X(t)\exp(-tX) =: r_t(X)$.

The IS algorithm does not change: Simulate independent realisations of X_i in $(\Omega, \mathcal{F}, Q_t)$ and set $\hat{\theta}_n^{(IS)} = (1/n)\sum_{i=1}^n X_i r_t(X_i)$.

IS in the case of Bernoulli mixture models

(see Glasserman and Li (2003))

Consider the loss function of a credit portfolio $L = \sum_{i=1}^{m} e_i Y_i$.

IS in the case of Bernoulli mixture models

(see Glasserman and Li (2003))
Consider the loss function of a credit portfolio $L = \sum_{i=1}^{m} e_i Y_i$.

Y_i are the loss indicators with default probability $\bar{p}_i$ and $e_i = (1 - \lambda_i) L_i$ are the positive deterministic exposures in the case that a corresponding loss happens. λ_i are the recovery rates and L_i are the credit nominals, for $i = 1, 2, \ldots, m$.

IS in the case of Bernoulli mixture models

(see Glasserman and Li (2003))
Consider the loss function of a credit portfolio $L = \sum_{i=1}^{m} e_i Y_i$.

Y_i are the loss indicators with default probability $\bar{p}_i$ and $e_i = (1 - \lambda_i) L_i$ are the positive deterministic exposures in the case that a corresponding loss happens. λ_i are the recovery rates and L_i are the credit nominals, for $i = 1, 2, \ldots, m$.

Let Z be a vector of economical impact factors, such that $Y_i|Z$ are independent and $Y_i|(Z = z) \sim \text{Bernoulli}(p_i(z))$, $\forall i = 1, 2, \ldots, m$.

IS in the case of Bernoulli mixture models

(see Glasserman and Li (2003))

Consider the loss function of a credit portfolio $L = \sum_{i=1}^{m} e_i Y_i$.

Y_i are the loss indicators with default probability $\bar{p}_i$ and $e_i = (1 - \lambda_i) L_i$ are the positive deterministic exposures in the case that a corresponding loss happens. λ_i are the recovery rates and L_i are the credit nominals, for $i = 1, 2, \ldots, m$.

Let Z be a vector of economical impact factors, such that $Y_i | Z$ are independent and $Y_i | (Z = z) \sim \mathrm{Bernoulli}(p_i(z))$, $\forall i = 1, 2, \ldots, m$.

Goal: Estimation of $\theta = P(L \geq c)$ by means of IS, for some given c with $c >> E(L)$.

IS in the case of Bernoulli mixture models

(see Glasserman and Li (2003))
Consider the loss function of a credit portfolio $L = \sum_{i=1}^{m} e_i Y_i$.

Y_i are the loss indicators with default probability $\bar{p}_i$ and $e_i = (1-\lambda_i)L_i$ are the positive deterministic exposures in the case that a corresponding loss happens. λ_i are the recovery rates and L_i are the credit nominals, for $i = 1, 2, \ldots, m$.

Let Z be a vector of economical impact factors, such that $Y_i|Z$ are independent and $Y_i|(Z = z) \sim \text{Bernoulli}(p_i(z))$, $\forall i = 1, 2, \ldots, m$.

Goal: Estimation of $\theta = P(L \geq c)$ by means of IS, for some given c with $c >> E(L)$.

Simplified case: Y_i are independent for $i = 1, 2, \ldots, m$.
Let $\Omega = \{0, 1\}^m$ be the state space of the random vector Y.
Consider the probability measure P in Ω:

$$P(\{y\}) = \prod_{i=1}^{m} \bar{p}_i^{y_i}(1-\bar{p}_i)^{1-y_i}, \; y \in \{0,1\}^m.$$

The moment generating function of L is $M_L(t) = \prod_{i=1}^{m}(e^{te_i}\bar{p}_i + 1 - \bar{p}_i)$.

IS in the case of Bernoulli mixture models (contd.)

Consider a probability measure Q_t:

$$Q_t(\{y\}) = \prod_{i=1}^{n} \left(\frac{\exp\{te_i y_i\}}{\exp\{te_i\}\bar{p}_i + 1 - \bar{p}_i} \bar{p}_i^{y_i} (1 - \bar{p}_i)^{1-y_i} \right).$$

IS in the case of Bernoulli mixture models (contd.)

Consider a probability measure Q_t:

$$Q_t(\{y\}) = \prod_{i=1}^{n} \left(\frac{\exp\{te_i y_i\}}{\exp\{te_i\}\bar{p}_i + 1 - \bar{p}_i} \bar{p}_i^{y_i} (1 - \bar{p}_i)^{1-y_i} \right).$$

Let $\bar{q}_{t,i}$ be new default probabilities

$$\bar{q}_{t,i} := \exp\{te_i\}\bar{p}_i / (\exp\{te_i\}\bar{p}_i + 1 - \bar{p}_i).$$

IS in the case of Bernoulli mixture models (contd.)

Consider a probability measure Q_t:

$$Q_t(\{y\}) = \prod_{i=1}^{n} \left(\frac{\exp\{te_i y_i\}}{\exp\{te_i\}\bar{p}_i + 1 - \bar{p}_i} \bar{p}_i^{y_i} (1 - \bar{p}_i)^{1-y_i} \right).$$

Let $\bar{q}_{t,i}$ be new default probabilities

$$\bar{q}_{t,i} := \exp\{te_i\}\bar{p}_i / (\exp\{te_i\}\bar{p}_i + 1 - \bar{p}_i).$$

We have $Q_t(\{y\}) = \prod_{i=1}^{m} \bar{q}_i^{y_i} (1 - \bar{q}_i)^{1-y_i}$, for $y \in \{0, 1\}^m$.

IS in the case of Bernoulli mixture models (contd.)

Consider a probability measure Q_t:

$$Q_t(\{y\}) = \prod_{i=1}^{n} \left(\frac{\exp\{te_i y_i\}}{\exp\{te_i\}\bar{p}_i + 1 - \bar{p}_i} \bar{p}_i^{y_i} (1 - \bar{p}_i)^{1-y_i} \right).$$

Let $\bar{q}_{t,i}$ be new default probabilities

$$\bar{q}_{t,i} := \exp\{te_i\}\bar{p}_i / (\exp\{te_i\}\bar{p}_i + 1 - \bar{p}_i).$$

We have $Q_t(\{y\}) = \prod_{i=1}^{m} \bar{q}_i^{y_i} (1 - \bar{q}_i)^{1-y_i}$, for $y \in \{0, 1\}^m$.
Thus after applying the exponential tilting the default indicators are independent with new default probabilities $\bar{q}_{t,i}$.

IS in the case of Bernoulli mixture models (contd.)

Consider a probability measure Q_t:

$$Q_t(\{y\}) = \prod_{i=1}^{n} \left(\frac{\exp\{te_i y_i\}}{\exp\{te_i\}\bar{p}_i + 1 - \bar{p}_i} \bar{p}_i^{y_i} (1 - \bar{p}_i)^{1-y_i} \right).$$

Let $\bar{q}_{t,i}$ be new default probabilities

$$\bar{q}_{t,i} := \exp\{te_i\}\bar{p}_i / (\exp\{te_i\}\bar{p}_i + 1 - \bar{p}_i).$$

We have $Q_t(\{y\}) = \prod_{i=1}^{m} \bar{q}_i^{y_i} (1 - \bar{q}_i)^{1-y_i}$, for $y \in \{0,1\}^m$.

Thus after applying the exponential tilting the default indicators are independent with new default probabilities $\bar{q}_{t,i}$.

$\lim_{t\to\infty} \bar{q}_{t,i} = 1$ and $\lim_{t\to-\infty} \bar{q}_{t,i} = 0$ imply that $E^{Q_t}(L)$ takes all values in $(0, \sum_{i=1}^{m} e_i)$ for $t \in \mathbb{R}$.

IS in the case of Bernoulli mixture models (contd.)

Consider a probability measure Q_t:

$$Q_t(\{y\}) = \prod_{i=1}^{n} \left(\frac{\exp\{te_i y_i\}}{\exp\{te_i\}\bar{p}_i + 1 - \bar{p}_i} \bar{p}_i^{y_i} (1 - \bar{p}_i)^{1-y_i} \right).$$

Let $\bar{q}_{t,i}$ be new default probabilities

$$\bar{q}_{t,i} := \exp\{te_i\}\bar{p}_i / (\exp\{te_i\}\bar{p}_i + 1 - \bar{p}_i).$$

We have $Q_t(\{y\}) = \prod_{i=1}^{m} \bar{q}_i^{y_i} (1 - \bar{q}_i)^{1-y_i}$, for $y \in \{0,1\}^m$.

Thus after applying the exponential tilting the default indicators are independent with new default probabilities $\bar{q}_{t,i}$.

$\lim_{t\to\infty} \bar{q}_{t,i} = 1$ and $\lim_{t\to-\infty} \bar{q}_{t,i} = 0$ imply that $E^{Q_t}(L)$ takes all values in $(0, \sum_{i=1}^{m} e_i)$ for $t \in \mathbb{R}$.

Choose t, such that $\sum_{i=1}^{m} e_i \bar{q}_{t,i} = c$.

IS in the case of Bernoulli mixture models (contd.)

The general case: Y_i are independent conditional on Z

IS in the case of Bernoulli mixture models (contd.)

The general case: Y_i are independent conditional on Z

1. Step: Estimation of the conditional excess probabilites $\theta(z) := P(L \geq c | Z = z)$ for a given realisation z of the economic factor Z, by means of the IS approach for the simplified case.

IS in the case of Bernoulli mixture models (contd.)

The general case: Y_i are independent conditional on Z

1. Step: Estimation of the conditional excess probabilites $\theta(z) := P(L \geq c | Z = z)$ for a given realisation z of the economic factor Z, by means of the IS approach for the simplified case.

Algorithm: IS for the conditional loss distribution

(1) For a given z compute the conditional default probabilities $p_i(z)$ (as in the simplified case) and solve the equation

$$\sum_{i=1}^{m} e_i \frac{\exp\{te_i\} p_i(z)}{\exp\{te_i\} p_i(z) + 1 - p_i(z)} = c \, .$$

The solution $t = t(c, z)$ specifies the correct *degree of tilting*.

IS in the case of Bernoulli mixture models (contd.)

The general case: Y_i are independent conditional on Z

1. Step: Estimation of the conditional excess probabilites $\theta(z) := P(L \geq c | Z = z)$ for a given realisation z of the economic factor Z, by means of the IS approach for the simplified case.

Algorithm: IS for the conditional loss distribution

(1) For a given z compute the conditional default probabilities $p_i(z)$ (as in the simplified case) and solve the equation

$$\sum_{i=1}^{m} e_i \frac{\exp\{te_i\} p_i(z)}{\exp\{te_i\} p_i(z) + 1 - p_i(z)} = c .$$

The solution $t = t(c, z)$ specifies the correct *degree of tilting*.

(2) Generate n_1 conditional realisations of the vector of default indicators $(Y_1, \ldots, Y_m)$, Y_i are simulated from $\mathrm{Bernoulli}(q_i)$, $i = 1, 2, \ldots, m$, with

$$q_i = \frac{\exp\{t(c, z) e_i\} p_i(z)}{\exp\{t(c, z) e_i\} p_i(z) + 1 - p_i(z)} .$$

The general case (contd.)

The general case (contd.)

(3) Let $M_L(t,z) := \prod[\exp\{t(c,z)e_i\}p_i(z) + 1 - p_i(z)]$ be the conditional moment generating function of L. Let $L^{(1)}$, $L^{(2)}$,...,$L^{(n_1)}$ be the n_1 conditional realisations of L for the n_1 simulated realisations of $Y_1, Y_2, \ldots, Y_m$. Compute the IS-estimator for the tail probability of the conditional loss distribution:

$$\hat{\theta}_{n_1}^{(IS)}(z) = M_L(t(c,z),z)\frac{1}{n_1}\sum_{j=1}^{n_1} I_{L^{(j)} \geq c} \exp\{-t(c,z)L^{(j)}\}L^{(j)}.$$

The general case (contd.)

(3) Let $M_L(t,z) := \prod[\exp\{t(c,z)e_i\}p_i(z) + 1 - p_i(z)]$ be the conditional moment generating function of L. Let $L^{(1)}$, $L^{(2)}$,...,$L^{(n_1)}$ be the n_1 conditional realisations of L for the n_1 simulated realisations of $Y_1, Y_2, \ldots, Y_m$. Compute the IS-estimator for the tail probability of the conditional loss distribution:

$$\hat{\theta}_{n_1}^{(IS)}(z) = M_L(t(c,z),z)\frac{1}{n_1}\sum_{j=1}^{n_1} I_{L^{(j)} \geq c} \exp\{-t(c,z)L^{(j)}\}L^{(j)}.$$

2. Step: Estimation of the unconditional excess probability $\theta = P(L \geq c)$.

The general case (contd.)

(3) Let $M_L(t,z) := \prod[\exp\{t(c,z)e_i\}p_i(z) + 1 - p_i(z)]$ be the conditional moment generating function of L. Let $L^{(1)}$, $L^{(2)}$,...,$L^{(n_1)}$ be the n_1 conditional realisations of L for the n_1 simulated realisations of $Y_1, Y_2, \ldots, Y_m$. Compute the IS-estimator for the tail probability of the conditional loss distribution:

$$\hat{\theta}_{n_1}^{(IS)}(z) = M_L(t(c,z),z)\frac{1}{n_1}\sum_{j=1}^{n_1} I_{L^{(j)} \geq c} \exp\{-t(c,z)L^{(j)}\}L^{(j)}.$$

2. Step: Estimation of the unconditional excess probability $\theta = P(L \geq c)$.

Naive approach: Generate many realisations z of the impact factors Z and compute $\hat{\theta}_{n_1}^{(IS)}(z)$ for every one of them. The required estimator is the average of $\hat{\theta}_{n_1}^{(IS)}(z)$ over all realisations z.
This is not the most efficient approach, see Glasserman and Li (2003).

The general case (contd.)

(3) Let $M_L(t,z) := \prod[\exp\{t(c,z)e_i\}p_i(z) + 1 - p_i(z)]$ be the conditional moment generating function of L. Let $L^{(1)}$, $L^{(2)}$,...,$L^{(n_1)}$ be the n_1 conditional realisations of L for the n_1 simulated realisations of $Y_1, Y_2, \ldots, Y_m$. Compute the IS-estimator for the tail probability of the conditional loss distribution:

$$\hat{\theta}_{n_1}^{(IS)}(z) = M_L(t(c,z),z)\frac{1}{n_1}\sum_{j=1}^{n_1} I_{L^{(j)} \geq c} \exp\{-t(c,z)L^{(j)}\}L^{(j)}.$$

2. Step: Estimation of the unconditional excess probability $\theta = P(L \geq c)$.

Naive approach: Generate many realisations z of the impact factors Z and compute $\hat{\theta}_{n_1}^{(IS)}(z)$ for every one of them. The required estimator is the average of $\hat{\theta}_{n_1}^{(IS)}(z)$ over all realisations z.
This is not the most efficient approach, see Glasserman and Li (2003).

A better alternative: IS for the impact factors.

IS for the impact factors

IS for the impact factors

Assumption: $Z \sim N_p(0, \Sigma)$ (e.g. probit-normal Bernoulli mixture)

IS for the impact factors

Assumption: $Z \sim N_p(0, \Sigma)$ (e.g. probit-normal Bernoulli mixture)

Let the IS density g be the density of $N_p(\mu, \Sigma)$ for a new expected vector $\mu \in \mathbb{R}^p$. A good choice of μ should lead to frequent realisations of z which imply high conditional default probabilities $p_i(z)$.

IS for the impact factors

Assumption: $Z \sim N_p(0, \Sigma)$ (e.g. probit-normal Bernoulli mixture)

Let the IS density g be the density of $N_p(\mu, \Sigma)$ for a new expected vector $\mu \in \mathbb{R}^p$. A good choice of μ should lead to frequent realisations of z which imply high conditional default probabilities $p_i(z)$.

The likelihood ratio:

$$r_\mu(Z) = \frac{\exp\{-\frac{1}{2}Z^t\Sigma^{-1}Z\}}{\exp\{-\frac{1}{2}(Z-\mu)^t\Sigma^{-1}(Z-\mu)\}} = \exp\{-\mu^t\Sigma^{-1}Z + \frac{1}{2}\mu^t\Sigma^1\mu\}$$

IS for the impact factors

Assumption: $Z \sim N_p(0, \Sigma)$ (e.g. probit-normal Bernoulli mixture)

Let the IS density g be the density of $N_p(\mu, \Sigma)$ for a new expected vector $\mu \in \mathbb{R}^p$. A good choice of μ should lead to frequent realisations of z which imply high conditional default probabilities $p_i(z)$.

The likelihood ratio:

$$r_\mu(Z) = \frac{\exp\{-\frac{1}{2}Z^t\Sigma^{-1}Z\}}{\exp\{-\frac{1}{2}(Z-\mu)^t\Sigma^{-1}(Z-\mu)\}} = \exp\{-\mu^t\Sigma^{-1}Z + \frac{1}{2}\mu^t\Sigma^1\mu\}$$

Algorithm: complete IS for Bernoulli mixture models with Gaussian factors

(1) Generate $z_1, z_2, \ldots, z_n \sim N_p(\mu, \Sigma)$ (n is the number of the simulation rounds)

IS for the impact factors

Assumption: $Z \sim N_p(0, \Sigma)$ (e.g. probit-normal Bernoulli mixture)

Let the IS density g be the density of $N_p(\mu, \Sigma)$ for a new expected vector $\mu \in \mathbb{R}^p$. A good choice of μ should lead to frequent realisations of z which imply high conditional default probabilities $p_i(z)$.

The likelihood ratio:

$$r_\mu(Z) = \frac{\exp\{-\frac{1}{2}Z^t\Sigma^{-1}Z\}}{\exp\{-\frac{1}{2}(Z-\mu)^t\Sigma^{-1}(Z-\mu)\}} = \exp\{-\mu^t\Sigma^{-1}Z + \frac{1}{2}\mu^t\Sigma^{1}\mu\}$$

Algorithm: complete IS for Bernoulli mixture models with Gaussian factors

(1) Generate $z_1, z_2, \ldots, z_n \sim N_p(\mu, \Sigma)$ (n is the number of the simulation rounds)

(2) For each z_i compute $\hat{\theta}_{n_1}^{(IS)}(z_i)$ by applying the IS algorithm for the conditional loss.

IS for the impact factors

Assumption: $Z \sim N_p(0, \Sigma)$ (e.g. probit-normal Bernoulli mixture)

Let the IS density g be the density of $N_p(\mu, \Sigma)$ for a new expected vector $\mu \in \mathbb{R}^p$. A good choice of μ should lead to frequent realisations of z which imply high conditional default probabilities $p_i(z)$.

The likelihood ratio:

$$r_\mu(Z) = \frac{\exp\{-\frac{1}{2} Z^t \Sigma^{-1} Z\}}{\exp\{-\frac{1}{2}(Z-\mu)^t \Sigma^{-1}(Z-\mu)\}} = \exp\{-\mu^t \Sigma^{-1} Z + \frac{1}{2}\mu^t \Sigma^{1} \mu\}$$

Algorithm: complete IS for Bernoulli mixture models with Gaussian factors

(1) Generate $z_1, z_2, \ldots, z_n \sim N_p(\mu, \Sigma)$ (n is the number of the simulation rounds)

(2) For each z_i compute $\hat{\theta}_{n_1}^{(IS)}(z_i)$ by applying the IS algorithm for the conditional loss.

(3) compute the IS estimator for the independent excess probability:

$$\hat{\theta}_n^{(IS)} = \frac{1}{n} \sum_{i=1}^{n} r_\mu(z_i) \hat{\theta}_{n_1}^{(IS)}(z_i)$$

The choice of μ

The choice of μ

μ should be chosen such that the variance of the estimator is small.

The choice of μ

μ should be chosen such that the variance of the estimator is small.

A sketch of the idea of Glasserman and Li (2003):

The choice of μ

μ should be chosen such that the variance of the estimator is small.

A sketch of the idea of Glasserman and Li (2003):

Since $\hat{\theta}_{n_1}^{(IS)}(z) \approx P(L \geq c|Z = z)$, search for an appropriate IS density for the function $z \mapsto P(L \geq c|Z = z)$.

The choice of μ

μ should be chosen such that the variance of the estimator is small.

A sketch of the idea of Glasserman and Li (2003):

Since $\hat{\theta}_{n_1}^{(IS)}(z) \approx P(L \geq c|Z = z)$, search for an appropriate IS density for the function $z \mapsto P(L \geq c|Z = z)$.

Approach:
a) the optimal IS denstity g^* is proportional to
$P(L \geq c|Z = z) \exp\{-\frac{1}{2} z^t \Sigma^{-1} z\}$.

The choice of μ

μ should be chosen such that the variance of the estimator is small.

A sketch of the idea of Glasserman and Li (2003):

Since $\hat{\theta}_{n_1}^{(IS)}(z) \approx P(L \geq c | Z = z)$, search for an appropriate IS density for the function $z \mapsto P(L \geq c | Z = z)$.

Approach:
a) the optimal IS denstity g^* is proportional to
$P(L \geq c | Z = z) \exp\{-\frac{1}{2} z^t \Sigma^{-1} z\}$.
b) use as IS density a multivariate normal distribution with the same mode as the optimal IS density g^*.

The choice of μ

μ should be chosen such that the variance of the estimator is small.

A sketch of the idea of Glasserman and Li (2003):

Since $\hat{\theta}^{(IS)}_{n_1}(z) \approx P(L \geq c|Z = z)$, search for an appropriate IS density for the function $z \mapsto P(L \geq c|Z = z)$.

Approach:
a) the optimal IS denstity g^* is proportional to
$P(L \geq c|Z = z)\exp\{-\frac{1}{2}z^t\Sigma^{-1}z\}$.
b) use as IS density a multivariate normal distribution with the same mode as the optimal IS density g^*.

The mode of a multivariate normal distribution $N_p(\mu, \Sigma)$ equals the expected vector μ, thus determining μ leads to the following optimization problem:
$\mu = \text{argmax}_z \left\{P(L \geq c|Z = z)\exp\{-\frac{1}{2}z^t\Sigma^{-1}z\}\right\}$.

The choice of μ

μ should be chosen such that the variance of the estimator is small.

A sketch of the idea of Glasserman and Li (2003):

Since $\hat{\theta}_{n_1}^{(IS)}(z) \approx P(L \geq c | Z = z)$, search for an appropriate IS density for the function $z \mapsto P(L \geq c | Z = z)$.

Approach:
a) the optimal IS denstity g^* is proportional to
$P(L \geq c | Z = z) \exp\{-\frac{1}{2} z^t \Sigma^{-1} z\}$.
b) use as IS density a multivariate normal distribution with the same mode as the optimal IS density g^*.

The mode of a multivariate normal distribution $N_p(\mu, \Sigma)$ equals the expected vector μ, thus determining μ leads to the following optimization problem:
$\mu = \text{argmax}_z \left\{P(L \geq c | Z = z) \exp\{-\frac{1}{2} z^t \Sigma^{-1} z\}\right\}$.
This problem is hard to solve exactly; in general $P(L \geq c | Z = z)$ is not available in analytical form.

The choice of μ

μ should be chosen such that the variance of the estimator is small.

A sketch of the idea of Glasserman and Li (2003):

Since $\hat{\theta}_{n_1}^{(IS)}(z) \approx P(L \geq c | Z = z)$, search for an appropriate IS density for the function $z \mapsto P(L \geq c | Z = z)$.

Approach:
a) the optimal IS denstity g^* is proportional to $P(L \geq c | Z = z) \exp\{-\frac{1}{2} z^t \Sigma^{-1} z\}$.
b) use as IS density a multivariate normal distribution with the same mode as the optimal IS density g^*.

The mode of a multivariate normal distribution $N_p(\mu, \Sigma)$ equals the expected vector μ, thus determining μ leads to the following optimization problem:
$\mu = \text{argmax}_z \left\{P(L \geq c | Z = z) \exp\{-\frac{1}{2} z^t \Sigma^{-1} z\}\right\}$.
This problem is hard to solve exactly; in general $P(L \geq c | Z = z)$ is not available in analytical form.

Glasserman und Li (2003) propose some solution approaches.