What is credit risk?

What is credit risk?

Citation from McNeil, Frey und Embrechts (2005):

Credit risk is the risk that the value of a portfolio changes due to unexpected changes in the credit quality of issuers or trading partners. This subsumes both losses due to defaults and losses caused by changes in credit quality such as the downgrading of a counterparty in an internal or external rating system.

What is credit risk?

Citation from McNeil, Frey und Embrechts (2005):

Credit risk is the risk that the value of a portfolio changes due to unexpected changes in the credit quality of issuers or trading partners. This subsumes both losses due to defaults and losses caused by changes in credit quality such as the downgrading of a counterparty in an internal or external rating system.

Examples of finance instruments affected by credit risk

- bond portfolios
- OTC ("over the counter") transactions
- trades with credit derivatives
- ...

A generic model of credit risk

A generic model of credit risk

Let P be a portfolio of n risky bonds of nominal value L_i, $i = 1, 2, \ldots, n$, respectively, set up at time T_0.

A generic model of credit risk

Let P be a portfolio of n risky bonds of nominal value L_i, $i = 1, 2, \ldots, n$, respectively, set up at time T_0.

p_i: the probability that (the issuer of) bond i defaults until time T, $T > T_0$

A generic model of credit risk

Let P be a portfolio of n risky bonds of nominal value L_i, $i = 1, 2, \ldots, n$, respectively, set up at time T_0.

p_i: the probability that (the issuer of) bond i defaults until time T, $T > T_0$

$\lambda_i \in [0, 1]$: "recovery rate" of bond i

$1 - \lambda_i$: percentage of lost value of bond i in case of default until time T

A generic model of credit risk

Let P be a portfolio of n risky bonds of nominal value L_i, $i = 1, 2, \ldots, n$, respectively, set up at time T_0.

p_i: the probability that (the issuer of) bond i defaults until time T, $T > T_0$
$\lambda_i \in [0, 1]$: "recovery rate" of bond i
$1 - \lambda_i$: percentage of lost value of bond i in case of default until time T

The loss given default for bond i at time T: $LGD_i = (1 - \lambda_i)L_i$

A generic model of credit risk

Let P be a portfolio of n risky bonds of nominal value L_i, $i = 1, 2, \ldots, n$, respectively, set up at time T_0.

p_i: the probability that (the issuer of) bond i defaults until time T, $T > T_0$
$\lambda_i \in [0, 1]$: "recovery rate" of bond i
$1 - \lambda_i$: percentage of lost value of bond i in case of default until time T

The loss given default for bond i at time T: $LGD_i = (1 - \lambda_i)L_i$

Model the default of bond i until time T by a Bernoulli distributed r.v. X_i with with $p_i = P(X_i = 1)$:

$$X_i = \begin{cases} 1 & \text{bond } i \text{ defaults} \\ 0 & \text{otherwise} \end{cases}$$

A generic model of credit risk

Let P be a portfolio of n risky bonds of nominal value L_i, $i = 1, 2, \ldots, n$, respectively, set up at time T_0.

p_i: the probability that (the issuer of) bond i defaults until time T, $T > T_0$
$\lambda_i \in [0, 1]$: "recovery rate" of bond i
$1 - \lambda_i$: percentage of lost value of bond i in case of default until time T

The loss given default for bond i at time T: $LGD_i = (1 - \lambda_i)L_i$

Model the default of bond i until time T by a Bernoulli distributed r.v. X_i with with $p_i = P(X_i = 1)$:

$$X_i = \begin{cases} 1 & \text{bond } i \text{ defaults} \\ 0 & \text{otherwise} \end{cases}$$

Total loss at time T: $L = \sum_{i=1}^{n} X_i \cdot LGD_i = \sum_{i=1}^{n} X_i(1 - \lambda_i)L_i$.

A generic model of credit risk

Let P be a portfolio of n risky bonds of nominal value L_i, $i = 1, 2, \ldots, n$, respectively, set up at time T_0.

p_i: the probability that (the issuer of) bond i defaults until time T, $T > T_0$

$\lambda_i \in [0, 1]$: "recovery rate" of bond i

$1 - \lambda_i$: percentage of lost value of bond i in case of default until time T

The loss given default for bond i at time T: $LGD_i = (1 - \lambda_i)L_i$

Model the default of bond i until time T by a Bernoulli distributed r.v. X_i with with $p_i = P(X_i = 1)$:

$$X_i = \begin{cases} 1 & \text{bond } i \text{ defaults} \\ 0 & \text{otherwise} \end{cases}$$

Total loss at time T: $L = \sum_{i=1}^n X_i \cdot LGD_i = \sum_{i=1}^n X_i(1 - \lambda_i)L_i$.

L is a r.v. and its distribution depends from the c.d.f. of $(X_1, \ldots, X_n, \lambda_1, \ldots, \lambda_n)^T$ ab.

The simplest model

The simplest model

- $L_i = L_1$, $\forall i$
- recovery rates are deterministic and $\lambda_i = \lambda_1$, $\forall i$
- X_i are i.i.d. with $p_i = p$ for all i, for some $p \in (0, 1)$.

The simplest model

- $L_i = L_1, \forall i$
- recovery rates are deterministic and $\lambda_i = \lambda_1, \forall i$
- X_i are i.i.d. with $p_i = p$ for all i, for some $p \in (0, 1)$.

Then we have $L = LGD_1 \cdot N$ with $N = \sum_{i=1}^{n} X_i \sim Binomial(n, p)$.

The simplest model

- $L_i = L_1, \forall i$
- recovery rates are deterministic and $\lambda_i = \lambda_1, \forall i$
- X_i are i.i.d. with $p_i = p$ for all i, for some $p \in (0, 1)$.

Then we have $L = LGD_1 \cdot N$ with $N = \sum_{i=1}^{n} X_i \sim Binomial(n, p)$.

Models with latent variables

The obligors (bonds) are partitioned into $m + 1$ homogeneous categories such that all obligors of a group have the same default probability.

The simplest model

- $L_i = L_1, \forall i$
- recovery rates are deterministic and $\lambda_i = \lambda_1, \forall i$
- X_i are i.i.d. with $p_i = p$ for all i, for some $p \in (0, 1)$.

Then we have $L = LGD_1 \cdot N$ with $N = \sum_{i=1}^{n} X_i \sim Binomial(n, p)$.

Models with latent variables
The obligors (bonds) are partitioned into $m + 1$ homogeneous categories such that all obligors of a group have the same default probability.

Historical data about the number of defaulting obligors of a certain category are used to obtain an estimator for the default probability of that category.

The simplest model

- $L_i = L_1, \forall i$
- recovery rates are deterministic and $\lambda_i = \lambda_1, \forall i$
- X_i are i.i.d. with $p_i = p$ for all i, for some $p \in (0, 1)$.

Then we have $L = LGD_1 \cdot N$ with $N = \sum_{i=1}^{n} X_i \sim Binomial(n, p)$.

Models with latent variables

The obligors (bonds) are partitioned into $m + 1$ homogeneous categories such that all obligors of a group have the same default probability.

Historical data about the number of defaulting obligors of a certain category are used to obtain an estimator for the default probability of that category.

$S = (S_1, S_2, \ldots, S_n)$, $S_i \in \{0, 1, \ldots, m\}$, is a status vector representing the category assignment; $S_i = j \in \{1, 2, \ldots, m\}$ means that obligor i belongs to category j

The simplest model

- $L_i = L_1, \forall i$
- recovery rates are deterministic and $\lambda_i = \lambda_1, \forall i$
- X_i are i.i.d. with $p_i = p$ for all i, for some $p \in (0,1)$.

Then we have $L = LGD_1 \cdot N$ with $N = \sum_{i=1}^{n} X_i \sim Binomial(n, p)$.

Models with latent variables

The obligors (bonds) are partitioned into $m+1$ homogeneous categories such that all obligors of a group have the same default probability.

Historical data about the number of defaulting obligors of a certain category are used to obtain an estimator for the default probability of that category.

$S = (S_1, S_2, \ldots, S_n)$, $S_i \in \{0, 1, \ldots, m\}$, is a status vector representing the category assignment; $S_i = j \in \{1, 2, \ldots, m\}$ means that obligor i belongs to category j (e.g. categories could be the rating classes).

The simplest model

- $L_i = L_1$, $\forall i$
- recovery rates are deterministic and $\lambda_i = \lambda_1$, $\forall i$
- X_i are i.i.d. with $p_i = p$ for all i, for some $p \in (0,1)$.

Then we have $L = LGD_1 \cdot N$ with $N = \sum_{i=1}^{n} X_i \sim Binomial(n,p)$.

Models with latent variables

The obligors (bonds) are partitioned into $m+1$ homogeneous categories such that all obligors of a group have the same default probability.

Historical data about the number of defaulting obligors of a certain category are used to obtain an estimator for the default probability of that category.

$S = (S_1, S_2, \ldots, S_n)$, $S_i \in \{0, 1, \ldots, m\}$, is a status vector representing the category assignment; $S_i = j \in \{1, 2, \ldots, m\}$ means that obligor i belongs to category j (e.g. categories could be the rating classes).

$S_i = 0$ corresponds to default.

The simplest model

- $L_i = L_1$, $\forall i$
- recovery rates are deterministic and $\lambda_i = \lambda_1$, $\forall i$
- X_i are i.i.d. with $p_i = p$ for all i, for some $p \in (0,1)$.

Then we have $L = LGD_1 \cdot N$ with $N = \sum_{i=1}^n X_i \sim Binomial(n,p)$.

Models with latent variables
The obligors (bonds) are partitioned into $m+1$ homogeneous categories such that all obligors of a group have the same default probability.

Historical data about the number of defaulting obligors of a certain category are used to obtain an estimator for the default probability of that category.

$S = (S_1, S_2, \ldots, S_n)$, $S_i \in \{0,1,\ldots,m\}$, is a status vector representing the category assignment; $S_i = j \in \{1,2,\ldots,m\}$ means that obligor i belongs to category j (e.g. categories could be the rating classes).

$S_i = 0$ corresponds to default.

Then we have $X_i = \begin{cases} 0 & S_i \neq 0 \\ 1 & S_i = 0 \end{cases}$

Models with latent variables (contd.)

Models with latent variables (contd.)

$S = (S_1, S_2, \ldots, S_n)^T$ is modelled by means of latent variables $Y = (Y_1, Y_2, \ldots, Y_n)^T$, e.g. Y_i could be the value of the assets of obligor i

Models with latent variables (contd.)

$S = (S_1, S_2, \ldots, S_n)^T$ is modelled by means of latent variables $Y = (Y_1, Y_2, \ldots, Y_n)^T$, e.g. Y_i could be the value of the assets of obligor i (firm value models).

Models with latent variables (contd.)

$S = (S_1, S_2, \ldots, S_n)^T$ is modelled by means of latent variables $Y = (Y_1, Y_2, \ldots, Y_n)^T$, e.g. Y_i could be the value of the assets of obligor i (firm value models).

Let d_{ij}, $i = 1, 2, \ldots, n$, $j = 0, 1, \ldots, m+1$ be threshold values such that $d_{i,0} = -\infty$ und $d_{i,m+1} = \infty$ and $S_i = j \Longleftrightarrow Y_i \in (d_{i,j}, d_{i,j+1}]$.

Models with latent variables (contd.)

$S = (S_1, S_2, \ldots, S_n)^T$ is modelled by means of latent variables $Y = (Y_1, Y_2, \ldots, Y_n)^T$, e.g. Y_i could be the value of the assets of obligor i (firm value models).

Let d_{ij}, $i = 1, 2, \ldots, n$, $j = 0, 1, \ldots, m+1$ be threshold values such that $d_{i,0} = -\infty$ und $d_{i,m+1} = \infty$ and $S_i = j \Longleftrightarrow Y_i \in (d_{i,j}, d_{i,j+1}]$.

Let F_i be the distribution function of Y_i. The probability of default for obligor i is $p_i = F_i(d_{i,1})$.

Models with latent variables (contd.)

$S=(S_1,S_2,\ldots,S_n)^T$ is modelled by means of latent variables $Y=(Y_1,Y_2,\ldots,Y_n)^T$, e.g. Y_i could be the value of the assets of obligor i (firm value models).

Let d_{ij}, $i=1,2,\ldots,n$, $j=0,1,\ldots,m+1$ be threshold values such that $d_{i,0}=-\infty$ und $d_{i,m+1}=\infty$ and $S_i=j\Longleftrightarrow Y_i\in(d_{i,j},d_{i,j+1}]$.

Let F_i be the distribution function of Y_i. The probability of default for obligor i is $p_i=F_i(d_{i,1})$.

The probability that the fisrt k obligors default:

$$p_{1,2,\ldots,k}:=P(Y_1\le d_{1,1},Y_2\le d_{2,1},\ldots,Y_k\le d_{k,1})$$

$$=C(F_1(d_{1,1}),F_2(d_{2,1}),\ldots,F_k(d_{k,1}),1,1,\ldots,1)=C(p_1,p_2,\ldots,p_k,1,\ldots,1)$$

Thus the totalt defalut probability depends essentially on the copula C of $(Y_1,Y_2,\ldots,Y_n)$.

The KMV model (see www.moodysanalytics.com)

The KMV model (see `www.moodysanalytics.com`)

The status variables $S = (S_1, S_2, \ldots, S_n)$ can only take two values 0 or 1, i.e. $m = 1$.

The KMV model (see `www.moodysanalytics.com`)

The status variables $S = (S_1, S_2, \ldots, S_n)$ can only take two values 0 or 1, i.e. $m = 1$.

The latent variables $Y = (Y_1, Y_2, \ldots, Y_n)^T$ depend on the value of the assets of the obligors as follows.

The KMV model (see `www.moodysanalytics.com`)

The status variables $S = (S_1, S_2, \dots, S_n)$ can only take two values 0 or 1, i.e. $m = 1$.

The latent variables $Y = (Y_1, Y_2, \dots, Y_n)^T$ depend on the value of the assets of the obligors as follows.

Merton's model

The balance sheet of each firm consists of assets and liabilities. The latter are devided in debt and equities.

The KMV model (see www.moodysanalytics.com)

The status variables $S = (S_1, S_2, \ldots, S_n)$ can only take two values 0 or 1, i.e. $m = 1$.

The latent variables $Y = (Y_1, Y_2, \ldots, Y_n)^T$ depend on the value of the assets of the obligors as follows.

Merton's model

The balance sheet of each firm consists of assets and liabilities. The latter are devided in debt and equities.

Notations:

$V_{A,i}(T)$: value of assets of firm i at time point T

$K_i := K_i(T)$: value of the debt of firm i at time point T

$V_{E,i}(T)$: value of equity of firm i at time point T

Assumption: future asset value is modelled by a geometric Brownian motion

The KMV model (contd.)

The KMV model (contd.)

$$V_{A,i}(T) = V_{A,i}(t)\exp\left\{\left(\mu_{A,i} - \frac{\sigma_{A,i}^2}{2}\right)(T-t) + \sigma_{A,i}\left(W_i(T) - W_i(t)\right)\right\},$$

The KMV model (contd.)

$$V_{A,i}(T) = V_{A,i}(t)\exp\left\{\left(\mu_{A,i} - \frac{\sigma_{A,i}^2}{2}\right)(T-t) + \sigma_{A,i}\left(W_i(T) - W_i(t)\right)\right\},$$

where

$\mu_{A,i}$ is the drift, $\sigma_{A,i}$ is the volatility and $(W_i(t)\colon 0 \leq t \leq T)$ is a standard Brownian motion (or equivalently a Wiener process).

The KMV model (contd.)

$$V_{A,i}(T) = V_{A,i}(t)\exp\left\{\left(\mu_{A,i} - \frac{\sigma^2_{A,i}}{2}\right)(T-t) + \sigma_{A,i}\left(W_i(T) - W_i(t)\right)\right\},$$

where

$\mu_{A,i}$ is the drift, $\sigma_{A,i}$ is the volatility and $(W_i(t)\colon 0 \leq t \leq T)$ is a standard Brownian motion (or equivalently a Wiener process).

Hence $(W_i(T) - W_i(t)) \sim N(0, T-t)$ and $\ln V_{A,i}(T) \sim N(\mu, \sigma^2)$ with

The KMV model (contd.)

$$V_{A,i}(T) = V_{A,i}(t) \exp\left\{\left(\mu_{A,i} - \frac{\sigma_{A,i}^2}{2}\right)(T-t) + \sigma_{A,i}\left(W_i(T) - W_i(t)\right)\right\},$$

where

$\mu_{A,i}$ is the drift, $\sigma_{A,i}$ is the volatility and $(W_i(t)\colon 0 \le t \le T)$ is a standard Brownian motion (or equivalently a Wiener process).

Hence $(W_i(T) - W_i(t)) \sim N(0, T-t)$ and $\ln V_{A,i}(T) \sim N(\mu, \sigma^2)$ with $\mu = \ln V_{A,i}(t) + \left(\mu_{A,i} - \frac{\sigma_{A,i}^2}{2}\right)(T-t)$ and $\sigma^2 = \sigma_{A,i}^2(T-t)$.

The KMV model (contd.)

$$V_{A,i}(T) = V_{A,i}(t)\exp\left\{\left(\mu_{A,i} - \frac{\sigma_{A,i}^2}{2}\right)(T-t) + \sigma_{A,i}\left(W_i(T) - W_i(t)\right)\right\},$$

where

$\mu_{A,i}$ is the drift, $\sigma_{A,i}$ is the volatility and $(W_i(t)\colon 0 \le t \le T)$ is a standard Brownian motion (or equivalently a Wiener process).

Hence $(W_i(T) - W_i(t)) \sim N(0, T-t)$ and $\ln V_{A,i}(T) \sim N(\mu, \sigma^2)$ with $\mu = \ln V_{A,i}(t) + \left(\mu_{A,i} - \frac{\sigma_{A,i}^2}{2}\right)(T-t)$ and $\sigma^2 = \sigma_{A,i}^2(T-t)$.

Further $X_i = I_{(-\infty,K_i)}(V_{A,i}(T))$ holds.

The KMV model (contd.)

$$V_{A,i}(T) = V_{A,i}(t)\exp\left\{\left(\mu_{A,i} - \frac{\sigma_{A,i}^2}{2}\right)(T-t) + \sigma_{A,i}\left(W_i(T) - W_i(t)\right)\right\},$$

where

$\mu_{A,i}$ is the drift, $\sigma_{A,i}$ is the volatility and $(W_i(t)\colon 0 \le t \le T)$ is a standard Brownian motion (or equivalently a Wiener process).

Hence $(W_i(T) - W_i(t)) \sim N(0, T-t)$ and $\ln V_{A,i}(T) \sim N(\mu, \sigma^2)$ with $\mu = \ln V_{A,i}(t) + \left(\mu_{A,i} - \frac{\sigma_{A,i}^2}{2}\right)(T-t)$ and $\sigma^2 = \sigma_{A,i}^2(T-t)$.

Further $X_i = I_{(-\infty,K_i)}(V_{A,i}(T))$ holds.

Set $Y_i = \frac{W_i(T)-W_i(t)}{\sqrt{T-t}} \sim N(0,1)$.

The KMV model (contd.)

$$V_{A,i}(T) = V_{A,i}(t)\exp\left\{\left(\mu_{A,i} - \frac{\sigma_{A,i}^2}{2}\right)(T-t) + \sigma_{A,i}\left(W_i(T) - W_i(t)\right)\right\},$$

where

$\mu_{A,i}$ is the drift, $\sigma_{A,i}$ is the volatility and $(W_i(t)\colon 0 \le t \le T)$ is a standard Brownian motion (or equivalently a Wiener process).

Hence $(W_i(T) - W_i(t)) \sim N(0, T-t)$ and $\ln V_{A,i}(T) \sim N(\mu, \sigma^2)$ with $\mu = \ln V_{A,i}(t) + \left(\mu_{A,i} - \frac{\sigma_{A,i}^2}{2}\right)(T-t)$ and $\sigma^2 = \sigma_{A,i}^2(T-t)$.

Further $X_i = I_{(-\infty,K_i)}(V_{A,i}(T))$ holds.

Set $Y_i = \frac{W_i(T)-W_i(t)}{\sqrt{T-t}} \sim N(0,1)$.

Then we get: $X_i = I_{(-\infty,K_i)}(V_{A,i}(T)) = I_{(-\infty,-DD_i)}(Y_i)$ where

$$DD_i = \frac{\ln V_{A,i}(t) - \ln K_i + (\mu_{A,i} - \frac{\sigma_{A,i}^2}{2})(T-t)}{\sigma_{A,i}\sqrt{T-t}}$$

The KMV model (contd.)

$$V_{A,i}(T) = V_{A,i}(t)\exp\left\{\left(\mu_{A,i} - \frac{\sigma_{A,i}^2}{2}\right)(T-t) + \sigma_{A,i}\left(W_i(T) - W_i(t)\right)\right\},$$

where

$\mu_{A,i}$ is the drift, $\sigma_{A,i}$ is the volatility and $(W_i(t)\colon 0 \le t \le T)$ is a standard Brownian motion (or equivalently a Wiener process).

Hence $(W_i(T) - W_i(t)) \sim N(0, T-t)$ and $\ln V_{A,i}(T) \sim N(\mu, \sigma^2)$ with $\mu = \ln V_{A,i}(t) + \left(\mu_{A,i} - \frac{\sigma_{A,i}^2}{2}\right)(T-t)$ and $\sigma^2 = \sigma_{A,i}^2(T-t)$.

Further $X_i = I_{(-\infty,K_i)}(V_{A,i}(T))$ holds.

Set $Y_i = \frac{W_i(T)-W_i(t)}{\sqrt{T-t}} \sim N(0,1)$.

Then we get: $X_i = I_{(-\infty,K_i)}(V_{A,i}(T)) = I_{(-\infty,-DD_i)}(Y_i)$ where

$$DD_i = \frac{\ln V_{A,i}(t) - \ln K_i + (\mu_{A,i} - \frac{\sigma_{A,i}^2}{2})(T-t)}{\sigma_{A,i}\sqrt{T-t}}$$

DD_i is called *distance-to-default*.

The KMV model (contd.)

The KMV model (contd.)

Computation of the “distance to default”

The KMV model (contd.)

Computation of the "distance to default"

$V_A, i(t)$, $\mu_{A,i}$ and $\sigma_{A,i}$ are needed.

The KMV model (contd.)

Computation of the "distance to default"

$V_A, i(t)$, $\mu_{A,i}$ and $\sigma_{A,i}$ are needed.
Difficulty: $V_{A,i}(t)$ can not be observed directly.

The KMV model (contd.)

Computation of the "distance to default"

$V_A, i(t)$, $\mu_{A,i}$ and $\sigma_{A,i}$ are needed.
Difficulty: $V_{A,i}(t)$ can not be observed directly.

However $V_{E,i}(t)$ can be observed by looking at the market stock prices.

The KMV model (contd.)

Computation of the "distance to default"

$V_A, i(t)$, $\mu_{A,i}$ and $\sigma_{A,i}$ are needed.
Difficulty: $V_{A,i}(t)$ can not be observed directly.

However $V_{E,i}(t)$ can be observed by looking at the market stock prices.

KMVs viewpoint: the equity holders have the right, but not the obligation, to pay off the holders of the other liabilities and take over the remaining assets of the firm.

The KMV model (contd.)

Computation of the "distance to default"

$V_A, i(t)$, $\mu_{A,i}$ and $\sigma_{A,i}$ are needed.
Difficulty: $V_{A,i}(t)$ can not be observed directly.

However $V_{E,i}(t)$ can be observed by looking at the market stock prices.

KMVs viewpoint: the equity holders have the right, but not the obligation, to pay off the holders of the other liabilities and take over the remaining assets of the firm.

This can be seen as a call option on the firms assets with a strike price equal to the book value of the firms liabilities.

The KMV model (contd.)

Computation of the "distance to default"

$V_A, i(t)$, $\mu_{A,i}$ and $\sigma_{A,i}$ are needed.
Difficulty: $V_{A,i}(t)$ can not be observed directly.

However $V_{E,i}(t)$ can be observed by looking at the market stock prices.

KMVs viewpoint: the equity holders have the right, but not the obligation, to pay off the holders of the other liabilities and take over the remaining assets of the firm.

This can be seen as a call option on the firms assets with a strike price equal to the book value of the firms liabilities.

Thus $V_{E,i}(T) = \max\{V_{A,i}(T) - K_i, 0\}$.

The KMV model (contd.)

Computation of the "distance to default"

$V_A, i(t)$, $\mu_{A,i}$ and $\sigma_{A,i}$ are needed.
Difficulty: $V_{A,i}(t)$ can not be observed directly.

However $V_{E,i}(t)$ can be observed by looking at the market stock prices.

KMVs viewpoint: the equity holders have the right, but not the obligation, to pay off the holders of the other liabilities and take over the remaining assets of the firm.

This can be seen as a call option on the firms assets with a strike price equal to the book value of the firms liabilities.

Thus $V_{E,i}(T) = \max\{V_{A,i}(T) - K_i, 0\}$.

The Black-Scholes formula implies (option price theory):

The KMV model (contd.)

Computation of the "distance to default"

$V_A, i(t)$, $\mu_{A,i}$ and $\sigma_{A,i}$ are needed.
Difficulty: $V_{A,i}(t)$ can not be observed directly.

However $V_{E,i}(t)$ can be observed by looking at the market stock prices.

KMVs viewpoint: the equity holders have the right, but not the obligation, to pay off the holders of the other liabilities and take over the remaining assets of the firm.

This can be seen as a call option on the firms assets with a strike price equal to the book value of the firms liabilities.

Thus $V_{E,i}(T) = \max\{V_{A,i}(T) - K_i, 0\}$.

The Black-Scholes formula implies (option price theory):

$V_{E,i}(t) = C(V_{A,i}(t), r, \sigma_{A,i}) = V_{A,i}(t)\phi(e_1) - K_i e^{-r(T-t)}\phi(e_2)$,

The KMV model (contd.)

Computation of the "distance to default"

$V_A, i(t)$, $\mu_{A,i}$ and $\sigma_{A,i}$ are needed.
Difficulty: $V_{A,i}(t)$ can not be observed directly.

However $V_{E,i}(t)$ can be observed by looking at the market stock prices.

KMVs viewpoint: the equity holders have the right, but not the obligation, to pay off the holders of the other liabilities and take over the remaining assets of the firm.

This can be seen as a call option on the firms assets with a strike price equal to the book value of the firms liabilities.

Thus $V_{E,i}(T) = \max\{V_{A,i}(T) - K_i, 0\}$.

The Black-Scholes formula implies (option price theory):

$V_{E,i}(t) = C(V_{A,i}(t), r, \sigma_{A,i}) = V_{A,i}(t)\phi(e_1) - K_i e^{-r(T-t)}\phi(e_2)$, where

$e_1 = \frac{\ln(V_{A,i}(t) - \ln K_i + (r + \sigma_{A,i}^2/2)(T-t)}{\sigma_{A,i}(T-t)}$, $e_2 = e_1 - \sigma_{A,i}(T-t)$,

ϕ is the the standard normal distribution function and r is the risk free interest rate.

Computation of the "distance to default" (contd.)

Computation of the "distance to default" (contd.)

The KMV model also postulates
$\sigma_{E,i} = g(V_{A,i}(t), \sigma_{A,i}, r)$, where g is some suitably selected proprietary function.

Computation of the "distance to default" (contd.)

The KMV model also postulates
$\sigma_{E,i} = g(V_{A,i}(t), \sigma_{A,i}, r)$, where g is some suitably selected proprietary function.

$V_{E,i}(t)$ and $\sigma_{E,i}$ are estimated based on historical data and the system of equalities below is solved w.r.t. $V_{A,i}(t)$ and $\sigma_{A,i}$:

Computation of the "distance to default" (contd.)

The KMV model also postulates
$\sigma_{E,i} = g(V_{A,i}(t), \sigma_{A,i}, r)$, where g is some suitably selected proprietary function.

$V_{E,i}(t)$ and $\sigma_{E,i}$ are estimated based on historical data and the system of equalities below is solved w.r.t. $V_{A,i}(t)$ and $\sigma_{A,i}$:

$V_{E,i}(t) = C(V_{A,i}(t), r, \sigma_{A,i})$
$\sigma_{E,i} = g(V_{A,i}(t), \sigma_{A,i}, r)$

Computation of the "distance to default" (contd.)

The KMV model also postulates
$\sigma_{E,i} = g(V_{A,i}(t), \sigma_{A,i}, r)$, where g is some suitably selected proprietary function.

$V_{E,i}(t)$ and $\sigma_{E,i}$ are estimated based on historical data and the system of equalities below is solved w.r.t. $V_{A,i}(t)$ and $\sigma_{A,i}$:

$V_{E,i}(t) = C(V_{A,i}(t), r, \sigma_{A,i})$
$\sigma_{E,i} = g(V_{A,i}(t), \sigma_{A,i}, r)$

The values obtained for $V_{A,i}(t)$ and $\sigma_{A,i}$ are used to compute DD_i:

$DD_i = \frac{\ln V_{A,i}(t) - \ln K_i + (\mu_{A,i} - \frac{\sigma^2_{A,i}}{2})(T-t)}{\sigma_{A,i}\sqrt{T-t}}$.

Computation of the "distance to default" (contd.)

The KMV model also postulates
$\sigma_{E,i} = g(V_{A,i}(t), \sigma_{A,i}, r)$, where g is some suitably selected proprietary function.

$V_{E,i}(t)$ and $\sigma_{E,i}$ are estimated based on historical data and the system of equalities below is solved w.r.t. $V_{A,i}(t)$ and $\sigma_{A,i}$:

$V_{E,i}(t) = C(V_{A,i}(t), r, \sigma_{A,i})$
$\sigma_{E,i} = g(V_{A,i}(t), \sigma_{A,i}, r)$

The values obtained for $V_{A,i}(t)$ and $\sigma_{A,i}$ are used to compute DD_i:

$DD_i = \frac{\ln V_{A,i}(t) - \ln K_i + (\mu_{A,i} - \frac{\sigma^2_{A,i}}{2})(T-t)}{\sigma_{A,i}\sqrt{T-t}}$.

Then $P(V_{A,i}(T) < K_i) = P(Y_i < -DD_i)$ and in the general setup of the latent variable model with $m = 1$ we have $d_{i1} = -DD_i$.

The expected default frequency (EDF)

The expected default frequency (EDF)

In the KMV model the default probability is not computed by setting $p_i := P(Y_i < -DD_i)$.

The expected default frequency (EDF)

In the KMV model the default probability is not computed by setting $p_i := P(Y_i < -DD_i)$.

Alternative: historical data are used to identify companies which at some stage in their history had the same distance to default DD_i.

The expected default frequency (EDF)

In the KMV model the default probability is not computed by setting $p_i := P(Y_i < -DD_i)$.

Alternative: historical data are used to identify companies which at some stage in their history had the same distance to default DD_i.

Then the observed default frequency is used as an estimator for the default probability p_i.

The expected default frequency (EDF)

In the KMV model the default probability is not computed by setting $p_i := P(Y_i < -DD_i)$.

Alternative: historical data are used to identify companies which at some stage in their history had the same distance to default DD_i.

Then the observed default frequency is used as an estimator for the default probability p_i. This estimator is called *expected default frequency, (EDF)*.

The expected default frequency (EDF)

In the KMV model the default probability is not computed by setting $p_i := P(Y_i < -DD_i)$.

Alternative: historical data are used to identify companies which at some stage in their history had the same distance to default DD_i.

Then the observed default frequency is used as an estimator for the default probability p_i. This estimator is called *expected default frequency, (EDF)*.

Summary of the univariate KMV model to compute the default probability of a company:

- Estimate the asset value $V_{A,i}$ and the volatilty $\sigma_{A,i}$ by using observations of the market value and the volatility of equity $V_{E,i}$, $\sigma_{E,i}$, the book of liabilities K_i, and by solving the system of equations above.
- Compute the distance-to-default DD_i by means of the corresponding formula.
- Estimate the default probability p_i in terms of the empirical distribution which relates the distance to default with the expected default frequency.

The multivariate KMV model: computation of multivariate default probabilities

The multivariate KMV model: computation of multivariate default probabilities

Let $Wj(t)$ be independent standard Brownian motions for $0 \leq t \leq T$, $j = 1, 2, \ldots, m$.

The multivariate KMV model: computation of multivariate default probabilities

Let $Wj(t)$ be independent standard Brownian motions for $0 \leq t \leq T$, $j = 1, 2, \ldots, m$.

Basic model: $V_{A,i}(T) =$

$$V_{A,i}(t) \exp \left\{ \left(\mu_{A,i} - \frac{\sigma_{A,i}^2}{2} \right) (T - t) + \sum_{j=1}^{m} \sigma_{A,i,j} \Bigg(W_j(T) - W_j(t) \Bigg) \right\},$$

where

The multivariate KMV model: computation of multivariate default probabilities

Let $Wj(t)$ be independent standard Brownian motions for $0 \le t \le T$, $j = 1, 2, \ldots, m$.

Basic model: $V_{A,i}(T) =$

$$V_{A,i}(t) \exp\left\{ \left(\mu_{A,i} - \frac{\sigma_{A,i}^2}{2}\right)(T - t) + \sum_{j=1}^{m} \sigma_{A,i,j} \Bigg(W_j(T) - W_j(t) \Bigg) \right\},$$

where

$\mu_{A,i}$ is the drift, $\sigma_{A,i}^2 = \sum_{j=1}^{m} \sigma_{A,i,j}^2$ is the volatility, and $\sigma_{A,i,j}$ quantifies the impact of the jth Brownian motion on the asset value of firm i.

The multivariate KMV model: computation of multivariate default probabilities

Let $Wj(t)$ be independent standard Brownian motions for $0 \leq t \leq T$, $j = 1, 2, \ldots, m$.

Basic model: $V_{A,i}(T) =$

$$V_{A,i}(t) \exp\left\{\left(\mu_{A,i} - \frac{\sigma_{A,i}^2}{2}\right)(T-t) + \sum_{j=1}^m \sigma_{A,i,j}\Bigg(W_j(T) - W_j(t)\Bigg)\right\},$$

where

$\mu_{A,i}$ is the drift, $\sigma_{A,i}^2 = \sum_{j=1}^m \sigma_{A,i,j}^2$ is the volatility, and $\sigma_{A,i,j}$ quantifies the impact of the jth Brownian motion on the asset value of firm i.

Set $Y_i := \frac{\sum_{j=1}^m \sigma_{A,i,j}(W_j(T) - W_j(t))}{\sigma_{A,i}\sqrt{T-t}}$. Then $Y = (Y_1, Y_2, \ldots, Y_n) \sim N(0, \Sigma)$,

The multivariate KMV model: computation of multivariate default probabilities

Let $Wj(t)$ be independent standard Brownian motions for $0 \leq t \leq T$, $j = 1, 2, \dots, m$.

Basic model: $V_{A,i}(T) =$

$$V_{A,i}(t) \exp \left\{ \left(\mu_{A,i} - \frac{\sigma_{A,i}^2}{2} \right) (T - t) + \sum_{j=1}^{m} \sigma_{A,i,j} \left(W_j(T) - W_j(t) \right) \right\},$$

where

$\mu_{A,i}$ is the drift, $\sigma_{A,i}^2 = \sum_{j=1}^{m} \sigma_{A,i,j}^2$ is the volatility, and $\sigma_{A,i,j}$ quantifies the impact of the jth Brownian motion on the asset value of firm i.

Set $Y_i := \frac{\sum_{j=1}^{m} \sigma_{A,i,j}(W_j(T) - W_j(t))}{\sigma_{A,i}\sqrt{T-t}}$. Then $Y = (Y_1, Y_2, \dots, Y_n) \sim N(0, \Sigma)$, where $\Sigma_{ij} = \frac{\sum_{k=1}^{m} \sigma_{A,i,k}\sigma_{A,j,k}}{\sigma_{A,i}\sigma_{A,j}}$.

The multivariate KMV model: computation of multivariate default probabilities

Let $Wj(t)$ be independent standard Brownian motions for $0 \leq t \leq T$, $j = 1, 2, \ldots, m$.

Basic model: $V_{A,i}(T) =$

$$V_{A,i}(t) \exp\left\{ \left(\mu_{A,i} - \frac{\sigma_{A,i}^2}{2}\right)(T - t) + \sum_{j=1}^{m} \sigma_{A,i,j} \Bigg(W_j(T) - W_j(t) \Bigg) \right\},$$

where

$\mu_{A,i}$ is the drift, $\sigma_{A,i}^2 = \sum_{j=1}^{m} \sigma_{A,i,j}^2$ is the volatility, and $\sigma_{A,i,j}$ quantifies the impact of the jth Brownian motion on the asset value of firm i.

Set $Y_i := \frac{\sum_{j=1}^{m} \sigma_{A,i,j}(W_j(T) - W_j(t))}{\sigma_{A,i}\sqrt{T-t}}$. Then $Y = (Y_1, Y_2, \ldots, Y_n) \sim N(0, \Sigma)$, where $\Sigma_{ij} = \frac{\sum_{k=1}^{m} \sigma_{A,i,k}\sigma_{A,j,k}}{\sigma_{A,i}\sigma_{A,j}}$.

We get $V_{A,i}(T) < K_i \Longleftrightarrow Y_i < -DD_i$ with

$$DD_i = \frac{\ln V_{A,i}(t) - \ln K_i + \left(\frac{-\sigma_{A,i}^2}{2} + \mu_{A,i} \right)(T-t)}{\sigma_{A,i}\sqrt{T-t}}.$$

The multivariate KMV model (contd.)

The multivariate KMV model (contd.)

The probability that the k first firms default:

$$P(X_1 = 1, X_2 = 1, \ldots, X_k = 1) = P(Y_1 < -DD_1, \ldots, Y_k < -DD_k)$$
$$= C_{\Sigma}^{Ga}(\phi(-DD_1), \ldots, \phi(-DD_k), 1, \ldots, 1),$$

The multivariate KMV model (contd.)

The probability that the k first firms default:

$$P(X_1 = 1, X_2 = 1, \ldots, X_k = 1) = P(Y_1 < -DD_1, \ldots, Y_k < -DD_k)$$
$$= C_{\Sigma}^{Ga}(\phi(-DD_1), \ldots, \phi(-DD_k), 1, \ldots, 1),$$

where C_{Σ}^{Ga} is the copula of a multivariate normal distribution with covariance matrix Σ.

The multivariate KMV model (contd.)

The probability that the k first firms default:

$$P(X_1 = 1, X_2 = 1, \dots, X_k = 1) = P(Y_1 < -DD_1, \dots, Y_k < -DD_k)$$
$$= C_{\Sigma}^{Ga}(\phi(-DD_1), \dots, \phi(-DD_k), 1, \dots, 1),$$

where C_{Σ}^{Ga} is the copula of a multivariate normal distribution with covariance matrix Σ.

Joint default frequency:

$$JDF_{1,2,\dots,k} = C_{\Sigma}^{Ga}(EDF_1, EDF_2, \dots, EDF_k, 1, \dots, 1),$$

where EDF_i is the default frequency for firm i, $i = 1, 2, \dots, k$.

Estimation of covariances/correlations $\sigma_{A,i,j}$

Estimation of covariances/correlations $\sigma_{A,i,j}$

Difficulties:

- n is typically quite large
- relatively few historical data available
- if n is large, then the pairwise estimated correlations coefficients do not build a positive correlation matrix, in general.

Estimation of covariances/correlations $\sigma_{A,i,j}$

Difficulties:

- n is typically quite large
- relatively few historical data available
- if n is large, then the pairwise estimated correlations coefficients do not build a positive correlation matrix, in general.

Possible approach:
Factor model for the latent variables in which the asset value of a company depends on certain common factors (macro-economical, global, regional, sector-based or country-based factors) and a company specific factor.

Estimation of covariances/correlations $\sigma_{A,i,j}$

Difficulties:

- n is typically quite large
- relatively few historical data available
- if n is large, then the pairwise estimated correlations coefficients do not build a positive correlation matrix, in general.

Possible approach:
Factor model for the latent variables in which the asset value of a company depends on certain common factors (macro-economical, global, regional, sector-based or country-based factors) and a company specific factor.

$Y = (Y_1, Y_2, \ldots, Y_n)^T = AZ + BU$ where
$Z = (Z_1, \ldots, Z_k)^T \sim N_k(0, \Lambda)$ are the k common factors,
$U = (U_1, \ldots, U_n)^T \sim N_n(0, I)$ are the company specific factors such that Z and U are independent, and the constant matrices $A = (a_{ij}) \in \mathbb{R}^{n \times k}$, $B = diag(b_1, \ldots, b_n) \in \mathbb{R}^{n \times n}$ are model parameters.

Estimation of covariances/correlations $\sigma_{A,i,j}$

Difficulties:

- n is typically quite large
- relatively few historical data available
- if n is large, then the pairwise estimated correlations coefficients do not build a positive correlation matrix, in general.

Possible approach:
Factor model for the latent variables in which the asset value of a company depends on certain common factors (macro-economical, global, regional, sector-based or country-based factors) and a company specific factor.

$Y = (Y_1, Y_2, \ldots, Y_n)^T = AZ + BU$ where
$Z = (Z_1, \ldots, Z_k)^T \sim N_k(0, \Lambda)$ are the k common factors,
$U = (U_1, \ldots, U_n)^T \sim N_n(0, I)$ are the company specific factors such that Z and U are independent, and the constant matrices $A = (a_{ij}) \in \mathbb{R}^{n\times k}$, $B = diag(b_1, \ldots, b_n) \in \mathbb{R}^{n\times n}$ are model parameters.

Then we have $cov(Y) = A\Lambda A^T + D$ where $D = diag(b_1^2, \ldots, b_n^2) \in \mathbb{R}^{n\times n}$.

Migration based models: Credit Metrics

Migration based models: Credit Metrics

It was developed by J.P.Morgan, see also MSCI (https://www.msci.com/)

Migration based models: Credit Metrics

It was developed by J.P.Morgan, see also MSCI (`https://www.msci.com/`)

It is primarily used fo the evaluation of bond portfolios (Siehe Crouhy et al. (2000)) and is based on a rating system (eg. *Moody's* or *Standard and Poor's*).

Migration based models: Credit Metrics

It was developed by J.P.Morgan, see also MSCI (`https://www.msci.com/`)

It is primarily used fo the evaluation of bond portfolios (Siehe Crouhy et al. (2000)) and is based on a rating system (eg. *Moody's* or *Standard and Poor's*).

It considers the changes of the portfolio value due to changes on the corresponding rating categories of the assets.

Migration based models: Credit Metrics

It was developed by J.P.Morgan, see also MSCI (`https://www.msci.com/`)

It is primarily used fo the evaluation of bond portfolios (Siehe Crouhy et al. (2000)) and is based on a rating system (eg. *Moody's* or *Standard and Poor's*).

It considers the changes of the portfolio value due to changes on the corresponding rating categories of the assets.

Let P be a portfolio consisting of n credits with a fixed holding duration (eg. 1 year). Let S_i be the status variable for debtor i, where the states are $0, 1, \ldots, m$ and $S_i = 0$ corresponds to default.

Migration based models: Credit Metrics

It was developed by J.P.Morgan, see also MSCI (`https://www.msci.com/`)

It is primarily used fo the evaluation of bond portfolios (Siehe Crouhy et al. (2000)) and is based on a rating system (eg. *Moody's* or *Standard and Poor's*).

It considers the changes of the portfolio value due to changes on the corresponding rating categories of the assets.

Let P be a portfolio consisting of n credits with a fixed holding duration (eg. 1 year). Let S_i be the status variable for debtor i, where the states are $0, 1, \ldots, m$ and $S_i = 0$ corresponds to default.

Example: Rating system of Standard and Poor's
$m = 7$; $S_i = 0$ means default; $S_i = 1$ or CCC; $S_i = 2$ or B; $S_i = 3$ or BB; $S_i = 4$ or BBB; $S_i = 5$ or A; $S_i = 6$ or AA; $S_i = 7$ or AAA.

Migration based models: Credit Metrics (contd.)

Migration based models: Credit Metrics (contd.)

For each debtor the dynamics of the status variable is modelled by means of a Markov chain with status set $\{0, 1, \ldots, m\}$ and transition matrix P.

Migration based models: Credit Metrics (contd.)

For each debtor the dynamics of the status variable is modelled by means of a Markov chain with status set $\{0, 1, \ldots, m\}$ and transition matrix P.

The transition probabilities are computed based on historical data: e.g.

Migration based models: Credit Metrics (contd.)

For each debtor the dynamics of the status variable is modelled by means of a Markov chain with status set $\{0, 1, \ldots, m\}$ and transition matrix P.

The transition probabilities are computed based on historical data: e.g.

Original state category	state category at the end of the year AAA	AA	A	BBB	BB	B	CCC	default
AAA	90.81	8.33	0.68	0.06	0.12	0	0	0
AA	0.70	90.65	7.79	0.64	0.06	0.14	0.02	0
A	0.09	2.27	91.05	5.52	0.74	0.26	0.01	0.06
BBB	0.02	0.33	5.95	86.93	5.30	1.17	0.12	0.18
BB	0.03	0.14	0.67	7.73	80.53	8.84	1.00	1.06
B	0	0.11	0.24	0.43	6.48	83.46	4.07	5.20
CCC	0.22	0	0.22	1.30	2.38	11.24	64.86	19.79

Migration based models: Credit Metrics (contd.)

For each debtor the dynamics of the status variable is modelled by means of a Markov chain with status set $\{0, 1, \dots, m\}$ and transition matrix P.

The transition probabilities are computed based on historical data: e.g.

Original state category	state category at the end of the year AAA	AA	A	BBB	BB	B	CCC	default
AAA	90.81	8.33	0.68	0.06	0.12	0	0	0
AA	0.70	90.65	7.79	0.64	0.06	0.14	0.02	0
A	0.09	2.27	91.05	5.52	0.74	0.26	0.01	0.06
BBB	0.02	0.33	5.95	86.93	5.30	1.17	0.12	0.18
BB	0.03	0.14	0.67	7.73	80.53	8.84	1.00	1.06
B	0	0.11	0.24	0.43	6.48	83.46	4.07	5.20
CCC	0.22	0	0.22	1.30	2.38	11.24	64.86	19.79

Recovery rates

In case of default the recovery rate depends on the status category of the defaulting debtor (prior to default). The mean and the standard deviation of the recovery rate are computed based on the historical data observed over time within each state category.

Evaluation of bonds if the status category changes

Evaluation of bonds if the status category changes

Example: Consider a BBB bond with maturity 5 years, a nominal value of 100 units and a coupon of 6% each year.
The forward *forward yield curves* for each status category are given as follows (in %):

Status	Year 1	Year 2	Year 3	Year 4
AAA	3.60	4.17	4.73	5.12
AA	3.65	4.22	4.78	5.17
A	3.73	4.32	4.93	5.32
BBB	4.10	4.67	5.25	5.63
BB	6.05	7.02	8.03	8.52
CCC	15.05	15.02	14.03	13.52

Evaluation of bonds if the status category changes

Example: Consider a BBB bond with maturity 5 years, a nominal value of 100 units and a coupon of 6% each year.
The forward *forward yield curves* for each status category are given as follows (in %):

Status	Year 1	Year 2	Year 3	Year 4
AAA	3.60	4.17	4.73	5.12
AA	3.65	4.22	4.78	5.17
A	3.73	4.32	4.93	5.32
BBB	4.10	4.67	5.25	5.63
BB	6.05	7.02	8.03	8.52
CCC	15.05	15.02	14.03	13.52

The bond pays 6 units at the end of the 4 years 1, 2, 3, 4 and 106 unit at the end of year 5.

Evaluation of bonds if the status category changes

Example: Consider a BBB bond with maturity 5 years, a nominal value of 100 units and a coupon of 6% each year.
The forward *forward yield curves* for each status category are given as follows (in %):

Status	Year 1	Year 2	Year 3	Year 4
AAA	3.60	4.17	4.73	5.12
AA	3.65	4.22	4.78	5.17
A	3.73	4.32	4.93	5.32
BBB	4.10	4.67	5.25	5.63
BB	6.05	7.02	8.03	8.52
CCC	15.05	15.02	14.03	13.52

The bond pays 6 units at the end of the 4 years 1, 2, 3, 4 and 106 unit at the end of year 5.

Assumption: At the end of the first year the bond is rated as an A bond. The value at the end of the first year:

$$V = 6 + \frac{6}{1+3,73\%} + \frac{6}{(1+4,32\%)^2} + \frac{6}{(1+4,93\%)^3} + \frac{106}{(1+5,32\%)^4} = 108.64$$

Evaluation of bonds if the status category changes (contd.)

Evaluation of bonds if the status category changes (contd.)

Example (contd.)

Analogous evaluation of the bond for other status category changes.

Evaluation of bonds if the status category changes (contd.)

Example (contd.)

Analogous evaluation of the bond for other status category changes.

Assumption: recovery rate in case of default is 51.13%.

Evaluation of bonds if the status category changes (contd.)

Example (contd.)

Analogous evaluation of the bond for other status category changes.

Assumption: recovery rate in case of default is 51.13%.

Status category at the end of the first year	value
AAA	109.35
AA	109.17
A	108.64
BBB	107.53
BB	102.01
B	98.09
CCC	83.63
Default	51.13

Use the transition probabilities of the Markov chain (estimated in terms of historical data) to compute the expected value of the bond at the end of the first year.

Value and risk of a bond portfolio in `Credit Metrics`

Value and risk of a bond portfolio in `Credit Metrics`

The return of bond i is modelled by a normal distribution Y_i.

Value and risk of a bond portfolio in `Credit Metrics`

The return of bond i is modelled by a normal distribution Y_i.

Let d_{Def}, d_{CCC}, ..., $d_{AAA} = +\infty$ be thresholds which define the transitions probabilities of debtor i at the end of the current period as follows:

$P(S_i = 0) = \phi(d_{Def})$, $P(S_i = CCC) = \phi(d_{CCC}) - \phi(d_{Def})$, ..., $P(S_i = AAA) = 1 - \phi(AA)$.

Value and risk of a bond portfolio in `Credit Metrics`

The return of bond i is modelled by a normal distribution Y_i.

Let d_{Def}, d_{CCC}, ..., $d_{AAA} = +\infty$ be thresholds which define the transitions probabilities of debtor i at the end of the current period as follows:

$P(S_i = 0) = \phi(d_{Def})$, $P(S_i = CCC) = \phi(d_{CCC}) - \phi(d_{Def})$, ..., $P(S_i = AAA) = 1 - \phi(AA)$.

The return of a vector of bonds is modelled as a multivariate normal distribution with correlation matrix R estimated by means of factor models.

Value and risk of a bond portfolio in `Credit Metrics`

The return of bond i is modelled by a normal distribution Y_i.

Let d_{Def}, d_{CCC}, ..., $d_{AAA} = +\infty$ be thresholds which define the transitions probabilities of debtor i at the end of the current period as follows:

$P(S_i = 0) = \phi(d_{Def})$, $P(S_i = CCC) = \phi(d_{CCC}) - \phi(d_{Def})$, ..., $P(S_i = AAA) = 1 - \phi(AA)$.

The return of a vector of bonds is modelled as a multivariate normal distribution with correlation matrix R estimated by means of factor models.

Joint probabilities of status category changes, e.g.

$$P(S_1 = 0, \ldots, S_n = 3) = P(Y_1 \leq d_{Def}, \ldots, d_B < Y_n \leq d_{BB})$$

can be then computed by using the Gaussian copula $C^{Ga}_{n,R}$ of $(Y_1, Y_2, \ldots, Y_n)$.

Value and risk of a bond portfolio in `Credit Metrics`

The return of bond i is modelled by a normal distribution Y_i.

Let d_{Def}, d_{CCC}, ..., $d_{AAA} = +\infty$ be thresholds which define the transitions probabilities of debtor i at the end of the current period as follows:

$P(S_i = 0) = \phi(d_{Def})$, $P(S_i = CCC) = \phi(d_{CCC}) - \phi(d_{Def})$, ..., $P(S_i = AAA) = 1 - \phi(AA)$.

The return of a vector of bonds is modelled as a multivariate normal distribution with correlation matrix R estimated by means of factor models.

Joint probabilities of status category changes, e.g.

$$P(S_1 = 0, \ldots, S_n = 3) = P(Y_1 \le d_{Def}, \ldots, d_B < Y_n \le d_{BB})$$

can be then computed by using the Gaussian copula $C^{Ga}_{n,R}$ of $(Y_1, Y_2, \ldots, Y_n)$.

Use simulation to compute the risk measures (VaR, CVaR) of the bond portfolio, e.g. by generating a large number of scenarios and then computing the empirical estimators of VaR, CVaR.

Approaches based on mixture models

Approaches based on mixture models

Assumptions:

(1) The default of each debtor depends on a number of (macro-economical) factors which are modelled stochastically.

Approaches based on mixture models

Assumptions:

(1) The default of each debtor depends on a number of (macro-economical) factors which are modelled stochastically.

(2) For a given realisation of these factors the defaults of different debtors are independent on each other.

Approaches based on mixture models

Assumptions:

(1) The default of each debtor depends on a number of (macro-economical) factors which are modelled stochastically.

(2) For a given realisation of these factors the defaults of different debtors are independent on each other.

Definition: The Bernoulli mixture distribution
The 0-1 random vector $X = (X_1, \ldots, X_n)^T$ has a *Bernoulli mixture distribution (BMD)* iff there exists a random vector $Z = (Z_1, Z_2, \ldots, Z_m)^T$, $m < n$, and the functions $f_i \colon \mathbb{R}^m \to [0, 1]$, $i = 1, 2, \ldots, n$, such that X conditioned on Z has independent components with $X_i | Z \sim \text{Bern}(f_i(Z))$.

Approaches based on mixture models

Assumptions:

(1) The default of each debtor depends on a number of (macro-economical) factors which are modelled stochastically.

(2) For a given realisation of these factors the defaults of different debtors are independent on each other.

Definition: The Bernoulli mixture distribution
The 0-1 random vector $X = (X_1, \ldots, X_n)^T$ has a *Bernoulli mixture distribution (BMD)* iff there exists a random vector $Z = (Z_1, Z_2, \ldots, Z_m)^T$, $m < n$, and the functions $f_i \colon \mathbb{R}^m \to [0, 1]$, $i = 1, 2, \ldots, n$, such that X conditioned on Z has independent components with $X_i|Z \sim \text{Bern}(f_i(Z))$.

Then $P(X = x|Z) = \prod_{i=1}^{n} f_i(Z)^{x_i}(1 - f_i(Z))^{1-x_i}$, $\forall x = (x_1, \ldots, x_n)^T \in \{0, 1\}^n$

Approaches based on mixture models

Assumptions:

(1) The default of each debtor depends on a number of (macro-economical) factors which are modelled stochastically.

(2) For a given realisation of these factors the defaults of different debtors are independent on each other.

Definition: The Bernoulli mixture distribution

The 0-1 random vector $X = (X_1, \ldots, X_n)^T$ has a *Bernoulli mixture distribution (BMD)* iff there exists a random vector $Z = (Z_1, Z_2, \ldots, Z_m)^T$, $m < n$, and the functions $f_i \colon \mathbb{R}^m \to [0,1]$, $i = 1, 2, \ldots, n$, such that X conditioned on Z has independent components with $X_i | Z \sim \text{Bern}(f_i(Z))$.

Then $P(X = x|Z) = \prod_{i=1}^n f_i(Z)^{x_i}(1 - f_i(Z))^{1-x_i}$, $\forall x = (x_1, \ldots, x_n)^T \in \{0,1\}^n$

The unconditional distribution:

$$P(X = x) = E(P(X = x|Z)) = E\left(\prod_{i=1}^n f_i(Z)^{x_i}(1 - f_i(Z))^{1-x_i}\right)$$

Approaches based on mixture models

Assumptions:

(1) The default of each debtor depends on a number of (macro-economical) factors which are modelled stochastically.

(2) For a given realisation of these factors the defaults of different debtors are independent on each other.

Definition: The Bernoulli mixture distribution

The 0-1 random vector $X = (X_1, \ldots, X_n)^T$ has a *Bernoulli mixture distribution (BMD)* iff there exists a random vector $Z = (Z_1, Z_2, \ldots, Z_m)^T$, $m < n$, and the functions $f_i \colon \mathbb{R}^m \to [0, 1]$, $i = 1, 2, \ldots, n$, such that X conditioned on Z has independent components with $X_i|Z \sim \text{Bern}(f_i(Z))$.

Then $P(X = x|Z) = \prod_{i=1}^n f_i(Z)^{x_i}(1 - f_i(Z))^{1-x_i}$, $\forall x = (x_1, \ldots, x_n)^T \in \{0, 1\}^n$

The unconditional distribution:

$$P(X = x) = E(P(X = x|Z)) = E\left(\prod_{i=1}^n f_i(Z)^{x_i}(1 - f_i(Z))^{1-x_i}\right)$$

If all function f_i coincide, i.e. $f_i = f$, $\forall i$, we get $N|Z \sim \text{Bin}(n, f(Z))$ for the number $N = \sum_{i=1}^n X_i$ of defaults.

The Poisson mixture distribution

The Poisson mixture distribution

Definition: The discrete random vector $X = (X_1, \ldots, X_n)^T$ has a *Poisson mixture distribution (PMD)* iff there exists a random vector $Z = (Z_1, Z_2, \ldots, Z_m)^T$, $m < n$, and the functions $\lambda_i \colon \mathbb{R}^m \to (0, \infty)$, $i = 1, 2, \ldots, n$, such that X conditioned on Z has independent components with $X_i | Z \sim Poi(\lambda_i(Z))$.

The Poisson mixture distribution

Definition: The discrete random vector $X = (X_1, \ldots, X_n)^T$ has a *Poisson mixture distribution (PMD)* iff there exists a random vector $Z = (Z_1, Z_2, \ldots, Z_m)^T$, $m < n$, and the functions $\lambda_i \colon \mathbb{R}^m \to (0, \infty)$, $i = 1, 2, \ldots, n$, such that X conditioned on Z has independent components with $X_i | Z \sim Poi(\lambda_i(Z))$.

Then $P(X = x | Z) = \prod_{i=1}^{n} \frac{\lambda_i(Z)^{x_i}}{x_i!} e^{-\lambda_i(Z)}$
$\forall x = (x_1, \ldots, x_n)^T \in (\mathbb{N} \cup \{0\})^n$.

The Poisson mixture distribution

Definition: The discrete random vector $X = (X_1, \dots, X_n)^T$ has a *Poisson mixture distribution (PMD)* iff there exists a random vector $Z = (Z_1, Z_2, \dots, Z_m)^T$, $m < n$, and the functions $\lambda_i \colon \mathbb{R}^m \to (0, \infty)$, $i = 1, 2, \dots, n$, such that X conditioned on Z has independent components with $X_i | Z \sim Poi(\lambda_i(Z))$.

Then $P(X = x | Z) = \prod_{i=1}^{n} \frac{\lambda_i(Z)^{x_i}}{x_i!} e^{-\lambda_i(Z)}$ $\forall x = (x_1, \dots, x_n)^T \in (\mathbb{N} \cup \{0\})^n$.

The unconditional distribution:

$$P(X = x) = E(P(X = x|Z)) = E\left(\prod_{i=1}^{n} \frac{\lambda_i(Z)^{x_i}}{x_i!} e^{-\lambda_i(Z)} \right)$$

The Poisson mixture distribution

Definition: The discrete random vector $X = (X_1, \ldots, X_n)^T$ has a *Poisson mixture distribution (PMD)* iff there exists a random vector $Z = (Z_1, Z_2, \ldots, Z_m)^T$, $m < n$, and the functions $\lambda_i \colon \mathbb{R}^m \to (0, \infty)$, $i = 1, 2, \ldots, n$, such that X conditioned on Z has independent components with $X_i | Z \sim Poi(\lambda_i(Z))$.

Then $P(X = x|Z) = \prod_{i=1}^n \frac{\lambda_i(Z)^{x_i}}{x_i!} e^{-\lambda_i(Z)}$
$\forall x = (x_1, \ldots, x_n)^T \in (\mathbb{N} \cup \{0\})^n$.

The unconditional distribution:

$$P(X = x) = E(P(X = x|Z)) = E\left(\prod_{i=1}^n \frac{\lambda_i(Z)^{x_i}}{x_i!} e^{-\lambda_i(Z)} \right)$$

Let $\bar{X}_i = I_{[1,\infty)}(X_i)$.
Then $\bar{X} = (\bar{X}_1, \ldots, \bar{X}_n)$ is BMD with $f_i(Z) = 1 - e^{-\lambda_i(Z)}$

The Poisson mixture distribution

Definition: The discrete random vector $X = (X_1, \ldots, X_n)^T$ has a *Poisson mixture distribution (PMD)* iff there exists a random vector $Z = (Z_1, Z_2, \ldots, Z_m)^T$, $m < n$, and the functions $\lambda_i : \mathbb{R}^m \to (0, \infty)$, $i = 1, 2, \ldots, n$, such that X conditioned on Z has independent components with $X_i|Z \sim Poi(\lambda_i(Z))$.

Then $P(X = x|Z) = \prod_{i=1}^{n} \frac{\lambda_i(Z)^{x_i}}{x_i!} e^{-\lambda_i(Z)}$
$\forall x = (x_1, \ldots, x_n)^T \in (\mathbb{N} \cup \{0\})^n$.

The unconditional distribution:

$$P(X = x) = E(P(X = x|Z)) = E\left(\prod_{i=1}^{n} \frac{\lambda_i(Z)^{x_i}}{x_i!} e^{-\lambda_i(Z)} \right)$$

Let $\bar{X}_i = I_{[1,\infty)}(X_i)$.
Then $\bar{X} = (\bar{X}_1, \ldots, \bar{X}_n)$ is BMD with $f_i(Z) = 1 - e^{-\lambda_i(Z)}$

If $\lambda_i(Z) << 1$ we get for the number $\tilde{N} = \sum_{i=1}^{n} \bar{X}_i \approx \sum_{i=1}^{n} X_i$ of defaults:

$$\tilde{N}|Z \sim Poisson(\bar{\lambda}(Z)), \text{ where } \bar{\lambda} = \sum_{i=1}^{n} \lambda_i(Z).$$

Examples of Bernoulli mixture distributions

Examples of Bernoulli mixture distributions

Assumptions :

- Z is univariate (i.e. there is only one risk factor)
- $f_i = f$, for all $i \in \{1, 2, \ldots, n\}$

Examples of Bernoulli mixture distributions

Assumptions :

- Z is univariate (i.e. there is only one risk factor)
- $f_i = f$, for all $i \in \{1, 2, \ldots, n\}$

We have $P(X_i = 1|Z) = f(Z)$, $\forall i$; $N|Z = \sum_{i=1}^{n} X_i \sim Bin(n, f(Z))$.

Examples of Bernoulli mixture distributions

Assumptions :

- Z is univariate (i.e. there is only one risk factor)
- $f_i = f$, for all $i \in \{1, 2, \ldots, n\}$

We have $P(X_i = 1|Z) = f(Z)$, $\forall i$; $N|Z = \sum_{i=1}^{n} X_i \sim Bin(n, f(Z))$.

The unconditional probability of default of the first k debtors is
$P(X_1 = 1, \ldots, X_k = 1, X_{k+1} = 0, \ldots, X_n = 0) =$
$E(P(X_1 = 1, \ldots, X_k = 1, X_{k+1} = 0, \ldots, X_n = 0|Z)) =$
$E(f(Z)^k (1 - f(Z))^{n-k})$

Examples of Bernoulli mixture distributions

Assumptions :

- Z is univariate (i.e. there is only one risk factor)
- $f_i = f$, for all $i \in \{1, 2, \ldots, n\}$

We have $P(X_i = 1|Z) = f(Z)$, $\forall i$; $N|Z = \sum_{i=1}^{n} X_i \sim Bin(n, f(Z))$.

The unconditional probability of default of the first k debtors is
$P(X_1 = 1, \ldots, X_k = 1, X_{k+1} = 0, \ldots, X_n = 0) =$
$E(P(X_1 = 1, \ldots, X_k = 1, X_{k+1} = 0, \ldots, X_n = 0|Z)) =$
$E(f(Z)^k(1 - f(Z))^{n-k})$

Let G be the distribution function of Z. Then
$P(X_1 = 1, \ldots, X_k = 1, X_{k+1} = 0, \ldots, X_n = 0) =$
$\int_{-\infty}^{\infty} f(z)^k(1 - f(z))^{n-k} d(G(z))$

Examples of Bernoulli mixture distributions

Assumptions :

- Z is univariate (i.e. there is only one risk factor)
- $f_i = f$, for all $i \in \{1, 2, \ldots, n\}$

We have $P(X_i = 1|Z) = f(Z)$, $\forall i$; $N|Z = \sum_{i=1}^{n} X_i \sim Bin(n, f(Z))$.

The unconditional probability of default of the first k debtors is
$P(X_1 = 1, \ldots, X_k = 1, X_{k+1} = 0, \ldots, X_n = 0) =$
$E(P(X_1 = 1, \ldots, X_k = 1, X_{k+1} = 0, \ldots, X_n = 0|Z)) =$
$E(f(Z)^k(1 - f(Z))^{n-k})$

Let G be the distribution function of Z. Then
$P(X_1 = 1, \ldots, X_k = 1, X_{k+1} = 0, \ldots, X_n = 0) =$
$\int_{-\infty}^{\infty} f(z)^k(1 - f(z))^{n-k} d(G(z))$

The distribution of the number N of defaults:

$$P(N = k) = \binom{n}{k} \int_{-\infty}^{\infty} f(z)^k(1 - f(z))^{n-k} d(G(z))$$

The beta-mixture distribution

The beta-mixture distribution

Let $Z \sim Beta(a, b)$ and $f(z) = z$.

The beta-mixture distribution

Let $Z \sim Beta(a, b)$ and $f(z) = z$.

The d.f. g of Z is given as $g(z) = \frac{1}{\beta(a,b)} z^{a-1}(1-z)^{b-1}$, for $a, b > 0$, $z \in (0, 1)$, where $\beta(a, b) = \int_0^1 z^{a-1}(1-z)^{b-1} dz$ is the Euler beta function.

The beta-mixture distribution

Let $Z \sim Beta(a, b)$ and $f(z) = z$.

The d.f. g of Z is given as $g(z) = \frac{1}{\beta(a,b)} z^{a-1}(1-z)^{b-1}$, for $a, b > 0$, $z \in (0,1)$, where $\beta(a,b) = \int_0^1 z^{a-1}(1-z)^{b-1}dz$ is the Euler beta function.

The distribution of the number of defaults:

$$P(N = k) = \binom{n}{k} \int_0^1 z^k (1-z)^{n-k} g(z)dz = \binom{n}{k} \frac{1}{\beta(a,b)} \int_0^1 z^{a+k-1}(1-z)^{n-k+b-1}dz$$

$$= \binom{n}{k} \frac{\beta(a+k, b+n-k)}{\beta(a,b)}$$ is the *beta-binomial distribution*

The beta-mixture distribution

Let $Z \sim Beta(a, b)$ and $f(z) = z$.

The d.f. g of Z is given as $g(z) = \frac{1}{\beta(a,b)} z^{a-1}(1-z)^{b-1}$, for $a, b > 0$, $z \in (0,1)$, where $\beta(a,b) = \int_0^1 z^{a-1}(1-z)^{b-1} dz$ is the Euler beta function.

The distribution of the number of defaults:

$$P(N = k) = \binom{n}{k} \int_0^1 z^k (1-z)^{n-k} g(z) dz = \binom{n}{k} \frac{1}{\beta(a,b)} \int_0^1 z^{a+k-1}(1-z)^{n-k+b-1} dz$$

$$= \binom{n}{k} \frac{\beta(a+k, b+n-k)}{\beta(a,b)}$$ is the *beta-binomial distribution*

The probit-normal mixture

The beta-mixture distribution

Let $Z \sim Beta(a,b)$ and $f(z) = z$.

The d.f. g of Z is given as $g(z) = \frac{1}{\beta(a,b)} z^{a-1}(1-z)^{b-1}$, for $a, b > 0$, $z \in (0,1)$, where $\beta(a,b) = \int_0^1 z^{a-1}(1-z)^{b-1} dz$ is the Euler beta function.

The distribution of the number of defaults:

$$P(N=k) = \binom{n}{k} \int_0^1 z^k (1-z)^{n-k} g(z) dz = \binom{n}{k} \frac{1}{\beta(a,b)} \int_0^1 z^{a+k-1}(1-z)^{n-k+b-1} dz$$

$$= \binom{n}{k} \frac{\beta(a+k, b+n-k)}{\beta(a,b)}$$ is the *beta-binomial distribution*

The probit-normal mixture

is obtained with $Z \sim N(0,1)$, $f(z) = \phi(\mu + \sigma z)$, $\mu \in \mathbb{R}$, $\sigma > 0$, where ϕ is the standard normal distribution.

The beta-mixture distribution

Let $Z \sim Beta(a,b)$ and $f(z) = z$.

The d.f. g of Z is given as $g(z) = \frac{1}{\beta(a,b)} z^{a-1}(1-z)^{b-1}$, for $a, b > 0$, $z \in (0,1)$, where $\beta(a,b) = \int_0^1 z^{a-1}(1-z)^{b-1}dz$ is the Euler beta function.

The distribution of the number of defaults:

$$P(N=k) = \binom{n}{k} \int_0^1 z^k (1-z)^{n-k} g(z) dz = \binom{n}{k} \frac{1}{\beta(a,b)} \int_0^1 z^{a+k-1}(1-z)^{n-k+b-1} dz$$

$$= \binom{n}{k} \frac{\beta(a+k, b+n-k)}{\beta(a,b)}$$ is the *beta-binomial distribution*

The probit-normal mixture

is obtained with $Z \sim N(0,1)$, $f(z) = \phi(\mu + \sigma z)$, $\mu \in \mathbb{R}$, $\sigma > 0$, where ϕ is the standard normal distribution.

The logit-normal mixture

The beta-mixture distribution

Let $Z \sim Beta(a, b)$ and $f(z) = z$.

The d.f. g of Z is given as $g(z) = \frac{1}{\beta(a,b)} z^{a-1}(1-z)^{b-1}$, for $a, b > 0$, $z \in (0, 1)$, where $\beta(a, b) = \int_0^1 z^{a-1}(1-z)^{b-1} dz$ is the Euler beta function.

The distribution of the number of defaults:

$$P(N = k) = \binom{n}{k} \int_0^1 z^k (1-z)^{n-k} g(z) dz = \binom{n}{k} \frac{1}{\beta(a, b)} \int_0^1 z^{a+k-1}(1-z)^{n-k+b-1} dz$$

$$= \binom{n}{k} \frac{\beta(a + k, b + n - k)}{\beta(a, b)}$$ is the *beta-binomial distribution*

The probit-normal mixture

is obtained with $Z \sim N(0, 1)$, $f(z) = \phi(\mu + \sigma z)$, $\mu \in \mathbb{R}$, $\sigma > 0$, where ϕ is the standard normal distribution.

The logit-normal mixture

is with $Z \sim N(0, 1)$, $f(z) = (1 + \exp\{\mu + \sigma z\})^{-1}$, $\mu \in \mathbb{R}$, $\sigma > 0$.

CreditRisk^{+} - a Poisson mixture model

CreditRisk$^+$ - a Poisson mixture model

(developed by CSFB in 1997, see Crouhy et al. (2000) and
`www.credit_suisse.com/investment_banking/research/en/credit_risk.jsp`

CreditRisk$^+$ - a Poisson mixture model

(developed by CSFB in 1997, see Crouhy et al. (2000) and
`www.credit_suisse.com/investment_banking/research/en/credit_risk.jsp`

Consider m independent risik factors $Z_1, Z_2, \ldots, Z_m$, $Z_j \sim \Gamma(\alpha_j, \beta_j)$, $j = 1, 2, \ldots, m$, with parameter α_j, β_j generally choosen such that such that $E(Z_j) = 1$.

CreditRisk$^+$ - a Poisson mixture model

(developed by CSFB in 1997, see Crouhy et al. (2000) and
`www.credit_suisse.com/investment_banking/research/en/credit_risk.jsp`

Consider m independent risik factors $Z_1, Z_2, \ldots, Z_m$, $Z_j \sim \Gamma(\alpha_j, \beta_j)$, $j = 1, 2, \ldots, m$, with parameter α_j, β_j generally choosen such that such that $E(Z_j) = 1$.

Let $\lambda_i(Z) = \bar{\lambda}_i \sum_{j=1}^{m} a_{ij} Z_j$, $\sum_{j=1}^{m} a_{ij} = 1$ for $i = 1, 2, \ldots, n$ for some parameters $\bar{\lambda}_i > 0$.

CreditRisk^{+} - a Poisson mixture model

(developed by CSFB in 1997, see Crouhy et al. (2000) and
`www.credit_suisse.com/investment_banking/research/en/credit_risk.jsp`

Consider m independent risik factors $Z_1, Z_2, \ldots, Z_m$, $Z_j \sim \Gamma(\alpha_j, \beta_j)$, $j = 1, 2, \ldots, m$, with parameter α_j, β_j generally choosen such that such that $E(Z_j) = 1$.

Let $\lambda_i(Z) = \bar{\lambda}_i \sum_{j=1}^{m} a_{ij} Z_j$, $\sum_{j=1}^{m} a_{ij} = 1$ for $i = 1, 2, \ldots, n$ for some parameters $\bar{\lambda}_i > 0$. Then $E(\lambda_i(Z)) = \bar{\lambda}_i > 0$) holds.

CreditRisk$^+$ - a Poisson mixture model

(developed by CSFB in 1997, see Crouhy et al. (2000) and
`www.credit_suisse.com/investment_banking/research/en/credit_risk.jsp`

Consider m independent risik factors $Z_1, Z_2, \ldots, Z_m$, $Z_j \sim \Gamma(\alpha_j, \beta_j)$, $j = 1, 2, \ldots, m$, with parameter α_j, β_j generally choosen such that such that $E(Z_j) = 1$.

Let $\lambda_i(Z) = \bar{\lambda}_i \sum_{j=1}^m a_{ij} Z_j$, $\sum_{j=1}^m a_{ij} = 1$ for $i = 1, 2, \ldots, n$ for some parameters $\bar{\lambda}_i > 0$. Then $E\left(\lambda_i(Z)\right) = \bar{\lambda}_i > 0)$ holds.

The density function of Z_j is given as $f_j(z) = \frac{z^{\alpha_j - 1} \exp\{-z/\beta_j\}}{\beta_j^{\alpha_j} \Gamma(\alpha_j)}$

CreditRisk$^+$ - a Poisson mixture model

(developed by CSFB in 1997, see Crouhy et al. (2000) and
www.credit_suisse.com/investment_banking/research/en/credit_risk.jsp

Consider m independent risik factors $Z_1, Z_2, \ldots, Z_m$, $Z_j \sim \Gamma(\alpha_j, \beta_j)$, $j = 1, 2, \ldots, m$, with parameter α_j, β_j generally choosen such that such that $E(Z_j) = 1$.

Let $\lambda_i(Z) = \bar{\lambda}_i \sum_{j=1}^m a_{ij} Z_j$, $\sum_{j=1}^m a_{ij} = 1$ for $i = 1, 2, \ldots, n$ for some parameters $\bar{\lambda}_i > 0$. Then $E\left(\lambda_i(Z)\right) = \bar{\lambda}_i > 0)$ holds.

The density function of Z_j is given as $f_j(z) = \frac{z^{\alpha_j - 1} \exp\{-z/\beta_j\}}{\beta_j^{\alpha_j} \Gamma(\alpha_j)}$

The loss given default for debtor i is $LGD_i = (1 - \lambda_i) L_i$, $1 \le i \le n$, where λ_i is the expected deterministic recovery rate and L_i is the amount of credit i.

CreditRisk$^+$ - a Poisson mixture model

(developed by CSFB in 1997, see Crouhy et al. (2000) and
www.credit_suisse.com/investment_banking/research/en/credit_risk.jsp

Consider m independent risik factors $Z_1, Z_2, \ldots, Z_m$, $Z_j \sim \Gamma(\alpha_j, \beta_j)$, $j = 1, 2, \ldots, m$, with parameter α_j, β_j generally choosen such that such that $E(Z_j) = 1$.

Let $\lambda_i(Z) = \bar{\lambda}_i \sum_{j=1}^{m} a_{ij} Z_j$, $\sum_{j=1}^{m} a_{ij} = 1$ for $i = 1, 2, \ldots, n$ for some parameters $\bar{\lambda}_i > 0$. Then $E(\lambda_i(Z)) = \bar{\lambda}_i > 0$) holds.

The density function of Z_j is given as $f_j(z) = \frac{z^{\alpha_j - 1} \exp\{-z/\beta_j\}}{\beta_j^{\alpha_j} \Gamma(\alpha_j)}$

The loss given default for debtor i is $LGD_i = (1 - \lambda_i) L_i$, $1 \le i \le n$, where λ_i is the expected deterministic recovery rate and L_i is the amount of credit i.

The goal: approximate the loss disribution by a discrete distribution durch and derive the generator function for the latter.

The probability generating function and its properties

The probability generating function and its properties

Let Y be a discrete r.v. taking values on $\{y_1, \ldots, y_m\}$ (a continuous r.v. with density function $f(y)$ in $\mathbb{R}$). The probability generating function (pgf) g_Y of Y is a mapping of $[0, 1]$ to the reals defined as

The probability generating function and its properties

Let Y be a discrete r.v. taking values on $\{y_1, \ldots, y_m\}$ (a continuous r.v. with density function $f(y)$ in $\mathbb{R}$). The probability generating function (pgf) g_Y of Y is a mapping of $[0,1]$ to the reals defined as $g_Y(t) := E(t^Y) = \sum_{i=1}^{m} t^{y_i} P(Y = y_i)$ $(g_Y(t) := \int_{-\infty}^{\infty} t^y f(y) dy)$.

The probability generating function and its properties

Let Y be a discrete r.v. taking values on $\{y_1, \ldots, y_m\}$ (a continuous r.v. with density function $f(y)$ in $\mathbb{R}$). The probability generating function (pgf) g_Y of Y is a mapping of $[0, 1]$ to the reals defined as $g_Y(t) := E(t^Y) = \sum_{i=1}^{m} t^{y_i} P(Y = y_i)$ $(g_Y(t) := \int_{-\infty}^{\infty} t^y f(y) dy)$.

Some properties of probability generating functions:

The probability generating function and its properties

Let Y be a discrete r.v. taking values on $\{y_1, \ldots, y_m\}$ (a continuous r.v. with density function $f(y)$ in $\mathbb{R}$). The probability generating function (pgf) g_Y of Y is a mapping of $[0, 1]$ to the reals defined as $g_Y(t) := E(t^Y) = \sum_{i=1}^m t^{y_i} P(Y = y_i)$ ($g_Y(t) := \int_{-\infty}^{\infty} t^y f(y) dy$).

Some properties of probability generating functions:

(i) If $Y \sim Bernoulli(p)$, then $g_Y(t) = 1 + p(t-1)$.

The probability generating function and its properties

Let Y be a discrete r.v. taking values on $\{y_1, \ldots, y_m\}$ (a continuous r.v. with density function $f(y)$ in $\mathbb{R}$). The probability generating function (pgf) g_Y of Y is a mapping of $[0, 1]$ to the reals defined as $g_Y(t) := E(t^Y) = \sum_{i=1}^{m} t^{y_i} P(Y = y_i)$ $(g_Y(t) := \int_{-\infty}^{\infty} t^y f(y) dy)$.

Some properties of probability generating functions:

(i) If $Y \sim Bernoulli(p)$, then $g_Y(t) = 1 + p(t - 1)$.

(ii) If $Y \sim Poisson(\lambda)$, then $g_Y(t) = \exp\{\lambda(t - 1)\}$.

The probability generating function and its properties

Let Y be a discrete r.v. taking values on $\{y_1, \ldots, y_m\}$ (a continuous r.v. with density function $f(y)$ in $\mathbb{R}$). The probability generating function (pgf) g_Y of Y is a mapping of $[0, 1]$ to the reals defined as $g_Y(t) := E(t^Y) = \sum_{i=1}^{m} t^{y_i} P(Y = y_i)$ $(g_Y(t) := \int_{-\infty}^{\infty} t^y f(y) dy)$.

Some properties of probability generating functions:

(i) If $Y \sim Bernoulli(p)$, then $g_Y(t) = 1 + p(t - 1)$.

(ii) If $Y \sim Poisson(\lambda)$, then $g_Y(t) = \exp\{\lambda(t - 1)\}$.

(iii) If the r.v. $X_1, \ldots, X_n$ are independent, then $g_{X_1+\ldots+X_n}(t) = \prod_{i=1}^{n} g_{X_i}(t)$.

The probability generating function and its properties

Let Y be a discrete r.v. taking values on $\{y_1, \ldots, y_m\}$ (a continuous r.v. with density function $f(y)$ in $\mathbb{R}$). The probability generating function (pgf) g_Y of Y is a mapping of $[0,1]$ to the reals defined as $g_Y(t) := E(t^Y) = \sum_{i=1}^m t^{y_i} P(Y = y_i)$ ($g_Y(t) := \int_{-\infty}^{\infty} t^y f(y) dy$).

Some properties of probability generating functions:

(i) If $Y \sim Bernoulli(p)$, then $g_Y(t) = 1 + p(t-1)$.

(ii) If $Y \sim Poisson(\lambda)$, then $g_Y(t) = \exp\{\lambda(t-1)\}$.

(iii) If the r.v. $X_1, \ldots, X_n$ are independent, then $g_{X_1+\ldots+X_n}(t) = \prod_{i=1}^n g_{X_i}(t)$.

(iv) Let Y be a r.v. with density function f and let $g_{X|Y=y}(t)$ be the pgf of $X|Y=y$. Then $g_X(t) = \int_{-\infty}^{\infty} g_{X|Y=y}(t) f(y) dy$.

The probability generating function and its properties

Let Y be a discrete r.v. taking values on $\{y_1, \ldots, y_m\}$ (a continuous r.v. with density function $f(y)$ in $\mathbb{R}$). The probability generating function (pgf) g_Y of Y is a mapping of $[0,1]$ to the reals defined as $g_Y(t) := E(t^Y) = \sum_{i=1}^m t^{y_i} P(Y = y_i)$ $(g_Y(t) := \int_{-\infty}^{\infty} t^y f(y) dy)$.

Some properties of probability generating functions:

(i) If $Y \sim \mathit{Bernoulli}(p)$, then $g_Y(t) = 1 + p(t-1)$.

(ii) If $Y \sim \mathit{Poisson}(\lambda)$, then $g_Y(t) = \exp\{\lambda(t-1)\}$.

(iii) If the r.v. $X_1, \ldots, X_n$ are independent, then $g_{X_1+\ldots+X_n}(t) = \prod_{i=1}^n g_{X_i}(t)$.

(iv) Let Y be a r.v. with density function f and let $g_{X|Y=y}(t)$ be the pgf of $X|Y=y$. Then $g_X(t) = \int_{-\infty}^{\infty} g_{X|Y=y}(t) f(y) dy$.

(v) Let $g_X(t)$ be the pgf of X. Then $P(X = k) = \frac{1}{k!} g_X^{(k)}(0)$, where $g_X^{(k)}(t) = \frac{d^k g_X(t)}{dt^k}$.

The pgf of the loss distribution

The pgf of the loss distribution

The loss will be approximated as an integer multiple of a prespecified loss unit L_0 (e.g. $L_o = 10^6$ Euro):

The pgf of the loss distribution

The loss will be approximated as an integer multiple of a prespecified loss unit L_0 (e.g. $L_o = 10^6$ Euro):

$LGD_i = (1-\lambda_i)L_i \approx \left[\frac{(1-\lambda_i)L_i}{L_0}\right] L_0 = v_i L_0$ with $v_i := \left[\frac{(1-\lambda_i)L_i}{L_0}\right]$,
where $[x] = \arg\min_t\{|t-x|\colon\, t \in \mathbb{Z}, t-x \in (-1/2, 1/2]\}$.

The pgf of the loss distribution

The loss will be approximated as an integer multiple of a prespecified loss unit L_0 (e.g. $L_o = 10^6$ Euro):

$LGD_i = (1-\lambda_i)L_i \approx \left[\frac{(1-\lambda_i)L_i}{L_0}\right] L_0 = v_i L_0$ with $v_i := \left[\frac{(1-\lambda_i)L_i}{L_0}\right]$,
where $[x] = \arg\min_t\{|t-x| \colon t \in \mathbb{Z}, t-x \in (-1/2, 1/2]\}$.

The loss function is then given by $L = \sum_{i=1}^n \bar{X}_i v_i L_0 \approx \sum_{i=1}^n X_i v_i L_0$, where $\bar{X}_i$ is the loss indicator and $(X_1, \ldots, X_n)$ has a PMD with factor vector $(Z_1, Z_2, \ldots, Z_m)$ as described above.

The pgf of the loss distribution

The loss will be approximated as an integer multiple of a prespecified loss unit L_0 (e.g. $L_o = 10^6$ Euro):

$LGD_i = (1-\lambda_i)L_i \approx \left[\frac{(1-\lambda_i)L_i}{L_0}\right] L_0 = v_i L_0$ with $v_i := \left[\frac{(1-\lambda_i)L_i}{L_0}\right]$,
where $[x] = \arg\min_t\{|t-x|: t \in \mathbb{Z}, t-x \in (-1/2, 1/2]\}$.

The loss function is then given by $L = \sum_{i=1}^n \bar{X}_i v_i L_0 \approx \sum_{i=1}^n X_i v_i L_0$, where $\bar{X}_i$ is the loss indicator and $(X_1, \ldots, X_n)$ has a PMD with factor vector $(Z_1, Z_2, \ldots, Z_m)$ as described above.

Step 1 Determine the pgf of (the approximative) number of losses $N = X_1 + \ldots + X_n$

The pgf of the loss distribution

The loss will be approximated as an integer multiple of a prespecified loss unit L_0 (e.g. $L_o = 10^6$ Euro):

$LGD_i = (1-\lambda_i)L_i \approx \left[\frac{(1-\lambda_i)L_i}{L_0}\right] L_0 = v_i L_0$ with $v_i := \left[\frac{(1-\lambda_i)L_i}{L_0}\right]$,
where $[x] = \arg\min_t\{|t-x|\colon t \in \mathbb{Z}, t-x \in (-1/2, 1/2]\}$.

The loss function is then given by $L = \sum_{i=1}^n \bar{X}_i v_i L_0 \approx \sum_{i=1}^n X_i v_i L_0$, where $\bar{X}_i$ is the loss indicator and $(X_1, \ldots, X_n)$ has a PMD with factor vector $(Z_1, Z_2, \ldots, Z_m)$ as described above.

Step 1 Determine the pgf of (the approximative) number of losses $N = X_1 + \ldots + X_n$

$X_i|Z \sim Poi(\lambda_i(Z)), \forall i \Longrightarrow g_{X_i|Z}(t) = \exp\{\lambda_i(Z)(t-1)\}, \forall i \Longrightarrow$
$g_{N|Z}(t) = \prod_{i=1}^n g_{X_i|Z}(t) = \prod_{i=1}^n \exp\{\lambda_i(Z)(t-1)\} = \exp\{\mu(t-1)\}$,
with $\mu := \sum_{i=1}^n \lambda_i(Z) = \sum_{i=1}^n \left(\bar{\lambda}_i \sum_{j=1}^m a_{ij} Z_j\right)$.

The pgf of the loss distribution (contd.)

The pgf of the loss distribution (contd.)

Then

$g_N(t) = \int_0^\infty \ldots \int_0^\infty g_{N|Z=(z_1,z_2,\ldots,z_m)} f_1(z_1) \ldots f_m(z_m) dz_1 \ldots dz_m =$

The pgf of the loss distribution (contd.)

Then

$g_N(t) = \int_0^\infty \dots \int_0^\infty g_{N|Z=(z_1,z_2,\dots,z_m)} f_1(z_1) \dots f_m(z_m) dz_1 \dots dz_m =$

$$\int_0^\infty \dots \int_0^\infty \exp\left\{ \sum_{i=1}^{n} \left(\bar{\lambda}_i \sum_{j=1}^{m} a_{ij} z_j \right) (t-1) \right\} f_1(z_1) \dots f_m(z_m) dz_1 \dots dz_m =$$

The pgf of the loss distribution (contd.)

Then

$g_N(t) = \int_0^\infty \dots \int_0^\infty g_{N|Z=(z_1,z_2,\dots,z_m)} f_1(z_1)\dots f_m(z_m) dz_1 \dots dz_m =$

$$\int_0^\infty \dots \int_0^\infty \exp\left\{ \sum_{i=1}^{n} \left(\bar{\lambda}_i \sum_{j=1}^{m} a_{ij} z_j \right) (t-1) \right\} f_1(z_1)\dots f_m(z_m) dz_1 \dots dz_m =$$

$$\int_0^\infty \dots \int_0^\infty \exp\left\{ (t-1) \sum_{j=1}^{m} \left(\underbrace{\sum_{i=1}^{n} \bar{\lambda}_i a_{ij}}_{\mu_j} \right) z_j) \right\} f_1(z_1)\dots f_m(z_m) dz_1 \dots dz_m =$$

The pgf of the loss distribution (contd.)

Then

$$g_N(t) = \int_0^\infty \dots \int_0^\infty g_{N|Z=(z_1,z_2,\dots,z_m)} f_1(z_1)\dots f_m(z_m) dz_1 \dots dz_m =$$

$$\int_0^\infty \dots \int_0^\infty \exp\left\{ \sum_{i=1}^n \left(\bar{\lambda}_i \sum_{j=1}^m a_{ij} z_j \right) (t-1) \right\} f_1(z_1)\dots f_m(z_m) dz_1 \dots dz_m =$$

$$\int_0^\infty \dots \int_0^\infty \exp\left\{ (t-1) \sum_{j=1}^m \left(\underbrace{\sum_{i=1}^n \bar{\lambda}_i a_{ij}}_{\mu_j} \right) z_j) \right\} f_1(z_1)\dots f_m(z_m) dz_1 \dots dz_m =$$

$$\int_0^\infty \dots \int_0^\infty \exp\{(t-1)\mu_1 z_1\} f_1(z_1) dz_1 \dots \exp\{(t-1)\mu_m z_m\} f_m(z_m) dz_m =$$

$$\prod_{j=1}^m \int_0^\infty \exp\{z_j \mu_j (t-1)\} \frac{1}{\beta_j^{\alpha_j} \Gamma(\alpha_j)} z_j^{\alpha_j - 1} \exp\{-z_j/\beta_j\} dz_j$$

The pgf of the loss distribution (contd.)

Then

$$g_N(t) = \int_0^\infty \dots \int_0^\infty g_{N|Z=(z_1,z_2,\dots,z_m)} f_1(z_1)\dots f_m(z_m) dz_1 \dots dz_m =$$

$$\int_0^\infty \dots \int_0^\infty \exp\left\{ \sum_{i=1}^{n} \left(\bar{\lambda}_i \sum_{j=1}^{m} a_{ij} z_j \right) (t-1) \right\} f_1(z_1)\dots f_m(z_m) dz_1 \dots dz_m =$$

$$\int_0^\infty \dots \int_0^\infty \exp\left\{ (t-1) \sum_{j=1}^{m} \left(\underbrace{\sum_{i=1}^{n} \bar{\lambda}_i a_{ij}}_{\mu_j} \right) z_j) \right\} f_1(z_1)\dots f_m(z_m) dz_1 \dots dz_m =$$

$$\int_0^\infty \dots \int_0^\infty \exp\{(t-1)\mu_1 z_1\} f_1(z_1) dz_1 \dots \exp\{(t-1)\mu_m z_m\} f_m(z_m) dz_m =$$

$$\prod_{j=1}^{m} \int_0^\infty \exp\{z_j \mu_j (t-1)\} \frac{1}{\beta_j^{\alpha_j} \Gamma(\alpha_j)} z_j^{\alpha_j - 1} \exp\{-z_j/\beta_j\} dz_j$$

The computation of each integral in the product obove yields

The pgf of the loss distribution (contd.)

Then

$g_N(t) = \int_0^\infty \dots \int_0^\infty g_{N|Z=(z_1,z_2,\dots,z_m)} f_1(z_1)\dots f_m(z_m)dz_1\dots dz_m =$

$$\int_0^\infty \dots \int_0^\infty \exp\left\{ \sum_{i=1}^n \left(\bar{\lambda}_i \sum_{j=1}^m a_{ij} z_j \right) (t-1) \right\} f_1(z_1)\dots f_m(z_m)dz_1\dots dz_m =$$

$$\int_0^\infty \dots \int_0^\infty \exp\left\{ (t-1) \sum_{j=1}^m \left(\underbrace{\sum_{i=1}^n \bar{\lambda}_i a_{ij}}_{\mu_j} \right) z_j) \right\} f_1(z_1)\dots f_m(z_m)dz_1\dots dz_m =$$

$$\int_0^\infty \dots \int_0^\infty \exp\{(t-1)\mu_1 z_1\} f_1(z_1)dz_1 \dots \exp\{(t-1)\mu_m z_m\} f_m(z_m)dz_m =$$

$$\prod_{j=1}^m \int_0^\infty \exp\{z_j\mu_j(t-1)\} \frac{1}{\beta_j^{\alpha_j}\Gamma(\alpha_j)} z_j^{\alpha_j-1} \exp\{-z_j/\beta_j\}dz_j$$

The computation of each integral in the product obove yields

$\int_0^\infty \frac{1}{\Gamma(\alpha_j)\beta_j^{\alpha_j}} \exp\{z_j\mu_j(t-1)\} z_j^{\alpha_j-1} \exp\{-z_j/\beta_j\}dz_j = \left(\frac{1-\delta_j}{1-\delta_j t}\right)^{\alpha_j}$ with $\delta_j = \beta_j\mu_j/(1+\beta_j\mu_j)$.

The pgf of the loss distribution (contd.)

The pgf of the loss distribution (contd.)

Thus we have $g_N(t) = \prod_{j=1}^{m} \left(\frac{1-\delta_j}{1-\delta_j t} \right)^{\alpha_j}$.

The pgf of the loss distribution (contd.)

Thus we have $g_N(t) = \prod_{j=1}^{m} \left(\frac{1-\delta_j}{1-\delta_j t} \right)^{\alpha_j}$.

Step 2 Determine the pgf of the (approximated) loss distribution $L = \sum_{i=1}^{n} X_i v_i L_0$.

The pgf of the loss distribution (contd.)

Thus we have $g_N(t) = \prod_{j=1}^{m} \left(\frac{1-\delta_j}{1-\delta_j t} \right)^{\alpha_j}$.

Step 2 Determine the pgf of the (approximated) loss distribution $L = \sum_{i=1}^{n} X_i v_i L_0$.

The conditional loss due to default of debtor i is $L_i | Z = v_i (X_i | Z)$

The pgf of the loss distribution (contd.)

Thus we have $g_N(t) = \prod_{j=1}^{m} \left(\frac{1-\delta_j}{1-\delta_j t} \right)^{\alpha_j}$.

Step 2 Determine the pgf of the (approximated) loss distribution $L = \sum_{i=1}^{n} X_i v_i L_0$.

The conditional loss due to default of debtor i is $L_i|Z = v_i(X_i|Z)$

$L_i|Z$ are independent for $i = 1, 2, \ldots, n \Longrightarrow$
$g_{L_i|Z}(t) = E(t^{L_i}|Z) = E(t^{v_i X_i}|Z) = g_{X_i|Z}(t^{v_i}) = \exp\{\lambda_i(Z)(t^{v_i} - 1)\}$.

The pgf of the loss distribution (contd.)

Thus we have $g_N(t) = \prod_{j=1}^{m} \left(\frac{1-\delta_j}{1-\delta_j t} \right)^{\alpha_j}$.

Step 2 Determine the pgf of the (approximated) loss distribution $L = \sum_{i=1}^{n} X_i v_i L_0$.

The conditional loss due to default of debtor i is $L_i | Z = v_i (X_i | Z)$

$L_i | Z$ are independent for $i = 1, 2, \ldots, n \Longrightarrow$
$g_{L_i|Z}(t) = E(t^{L_i}|Z) = E(t^{v_i X_i}|Z) = g_{X_i|Z}(t^{v_i}) = \exp\{\lambda_i(Z)(t^{v_i} - 1)\}$.
The pgf od the conditional overall loss is
$g_{L|Z}(t) = g_{L_1+L_2+\ldots+L_n|Z}(t) = \prod_{i=1}^{n} g_{L_i|Z}(t) =$
$\prod_{i=1}^{n} g_{X_i|Z}(t^{v_i}) = \exp\left\{ \sum_{j=1}^{m} Z_j \left(\sum_{i=1}^{n} \bar{\lambda}_i a_{ij} (t^{v_i} - 1) \right) \right\}$.

The pgf of the loss distribution (contd.)

Thus we have $g_N(t) = \prod_{j=1}^{m} \left(\frac{1-\delta_j}{1-\delta_j t} \right)^{\alpha_j}$.

Step 2 Determine the pgf of the (approximated) loss distribution $L = \sum_{i=1}^{n} X_i v_i L_0$.

The conditional loss due to default of debtor i is $L_i|Z = v_i(X_i|Z)$

$L_i|Z$ are independent for $i = 1, 2, \ldots, n \Longrightarrow$

$g_{L_i|Z}(t) = E(t^{L_i}|Z) = E(t^{v_i X_i}|Z) = g_{X_i|Z}(t^{v_i}) = \exp\{\lambda_i(Z)(t^{v_i} - 1)\}$.

The pgf od the conditional overall loss is

$g_{L|Z}(t) = g_{L_1+L_2+\ldots+L_n|Z}(t) = \prod_{i=1}^{n} g_{L_i|Z}(t) =$
$\prod_{i=1}^{n} g_{X_i|Z}(t^{v_i}) = \exp\left\{ \sum_{j=1}^{m} Z_j \left(\sum_{i=1}^{n} \bar{\lambda}_i a_{ij}(t^{v_i} - 1) \right) \right\}$.

Analogous computations as in the case of $g_N(t)$ yield:

$$g_L(t) = \prod_{j=1}^{m} \left(\frac{1-\delta_j}{1-\delta_j \Lambda_j(t)} \right)^{\alpha_j} \text{ wobei } \Lambda_j(t) = \frac{1}{\mu_j} \sum_{i=1}^{n} \bar{\lambda}_i a_{ij} t^{v_i}.$$

The pgf of the loss distribution (contd.)

The pgf of the loss distribution (contd.)

Example: Consider a credit portfolio with $n = 100$ credits, and m risk factors, where $m = 1$ or $m = 5$.

The pgf of the loss distribution (contd.)

Example: Consider a credit portfolio with $n = 100$ credits, and m risk factors, where $m = 1$ or $m = 5$.
Assume that $\bar{\lambda}_i = \bar{\lambda} = 0.15$, for $i = 1, 2, \ldots, n$, $\alpha_j = \alpha = 1$, $\beta_j = \beta = 1$, $a_{i,j} = 1/m$, $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, m$.

The pgf of the loss distribution (contd.)

Example: Consider a credit portfolio with $n = 100$ credits, and m risk factors, where $m = 1$ or $m = 5$.
Assume that $\bar{\lambda}_i = \bar{\lambda} = 0.15$, for $i = 1, 2, \ldots, n$, $\alpha_j = \alpha = 1$, $\beta_j = \beta = 1$, $a_{i,j} = 1/m$, $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, m$.

The probability that k creditors will default is given as follows for any $k \in \mathbb{N} \cup \{0\}$:

The pgf of the loss distribution (contd.)

Example: Consider a credit portfolio with $n = 100$ credits, and m risk factors, where $m = 1$ or $m = 5$.
Assume that $\bar{\lambda}_i = \bar{\lambda} = 0.15$, for $i = 1, 2, \ldots, n$, $\alpha_j = \alpha = 1$, $\beta_j = \beta = 1$, $a_{i,j} = 1/m$, $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, m$.

The probability that k creditors will default is given as follows for any $k \in \mathbb{N} \cup \{0\}$:
$P(N = k) = \frac{1}{k!} g_N^{(k)}(0) = \frac{1}{k!} \frac{d^k g_N}{dt^k}$.

The pgf of the loss distribution (contd.)

Example: Consider a credit portfolio with $n = 100$ credits, and m risk factors, where $m = 1$ or $m = 5$.
Assume that $\bar{\lambda}_i = \bar{\lambda} = 0.15$, for $i = 1, 2, \ldots, n$, $\alpha_j = \alpha = 1$, $\beta_j = \beta = 1$, $a_{i,j} = 1/m$, $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, m$.

The probability that k creditors will default is given as follows for any $k \in \mathbb{N} \cup \{0\}$:
$P(N = k) = \frac{1}{k!} g_N^{(k)}(0) = \frac{1}{k!} \frac{d^k g_N}{dt^k}$.

For the computation of $P(N = k)$, $k = 0, 1, \ldots, 100$, we can use the following recursive formula

The pgf of the loss distribution (contd.)

Example: Consider a credit portfolio with $n = 100$ credits, and m risk factors, where $m = 1$ or $m = 5$.
Assume that $\bar{\lambda}_i = \bar{\lambda} = 0.15$, for $i = 1, 2, \ldots, n$, $\alpha_j = \alpha = 1$, $\beta_j = \beta = 1$, $a_{i,j} = 1/m$, $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, m$.

The probability that k creditors will default is given as follows for any $k \in \mathbb{N} \cup \{0\}$:
$P(N = k) = \frac{1}{k!} g_N^{(k)}(0) = \frac{1}{k!} \frac{d^k g_N}{dt^k}$.

For the computation of $P(N = k)$, $k = 0, 1, \ldots, 100$, we can use the following recursive formula

$g_N^{(k)}(0) = \sum_{l=0}^{k-1} \binom{k-1}{l} g_N^{(k-1-l)}(0) \sum_{j=1}^{m} l! \alpha_j \delta_j^{l+1}$, where $k > 1$.

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

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 $L > q_\alpha(L)$.

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 $VaR_\alpha(L) = q_\alpha(L)$, $CVaR_\alpha = E(L|L > q_\alpha(L))$, $CVaR_{i,\alpha} = E(L_i|L > q_\alpha(L))$, 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.

$\widehat{CVaR}_\alpha^{(MC)}(L)$ 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 $(\Omega, \mathcal{F}, P)$ 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.

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,..., 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,. . . , 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 $\lim\limits_{n\to\infty} \hat{\theta}_n^{(MC)} = \theta$ 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,. . . , 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.

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

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

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

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

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

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 $t = argmin\{M_X(t)e^{-tc}\colon t \geq 0\}$ 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 $M_x(t)\colon \mathbb{R} \to \mathbb{R}$ 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 $t = argmin\{M_X(t) e^{-tc} \colon t \geq 0\}$ which imples $t = t(c)$, where $t(c)$ is the solution of the equation $\mu_t = c$.
(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:
$P(A) := \int_{x\in A} f(x)dx$ 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 $(\Omega, \mathcal{F}, P)$.

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.

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 $M_X(t) = E^P(\exp\{tX\}) < \infty$, $\forall t$.

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. $Q_t(A) := E^P\left(\frac{\exp\{tX\}}{M_X(t)}; A\right)$.

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 $Y_i|(Z = z) \sim Bernoulli(p_i(z))$, $\forall i = 1, 2, \ldots, m$.

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 $Y_i | (Z = z) \sim Bernoulli(p_i(z))$, $\forall i = 1, 2, \ldots, m$.

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 $Y_i | (Z = z) \sim Bernoulli(p_i(z))$, $\forall i = 1, 2, \ldots, m$.

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 $Bernoulli(q_i)$, $i = 1, 2, \ldots, m$, 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 $\hat{\theta}_{n_1}^{(IS)}(z) \approx P(L \geq c|Z = z)$, 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:
$\mu = \text{argmax}_z \left\{P(L \geq c|Z = z)\exp\{-\frac{1}{2}z^t\Sigma^{-1}z\}\right\}$.

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.