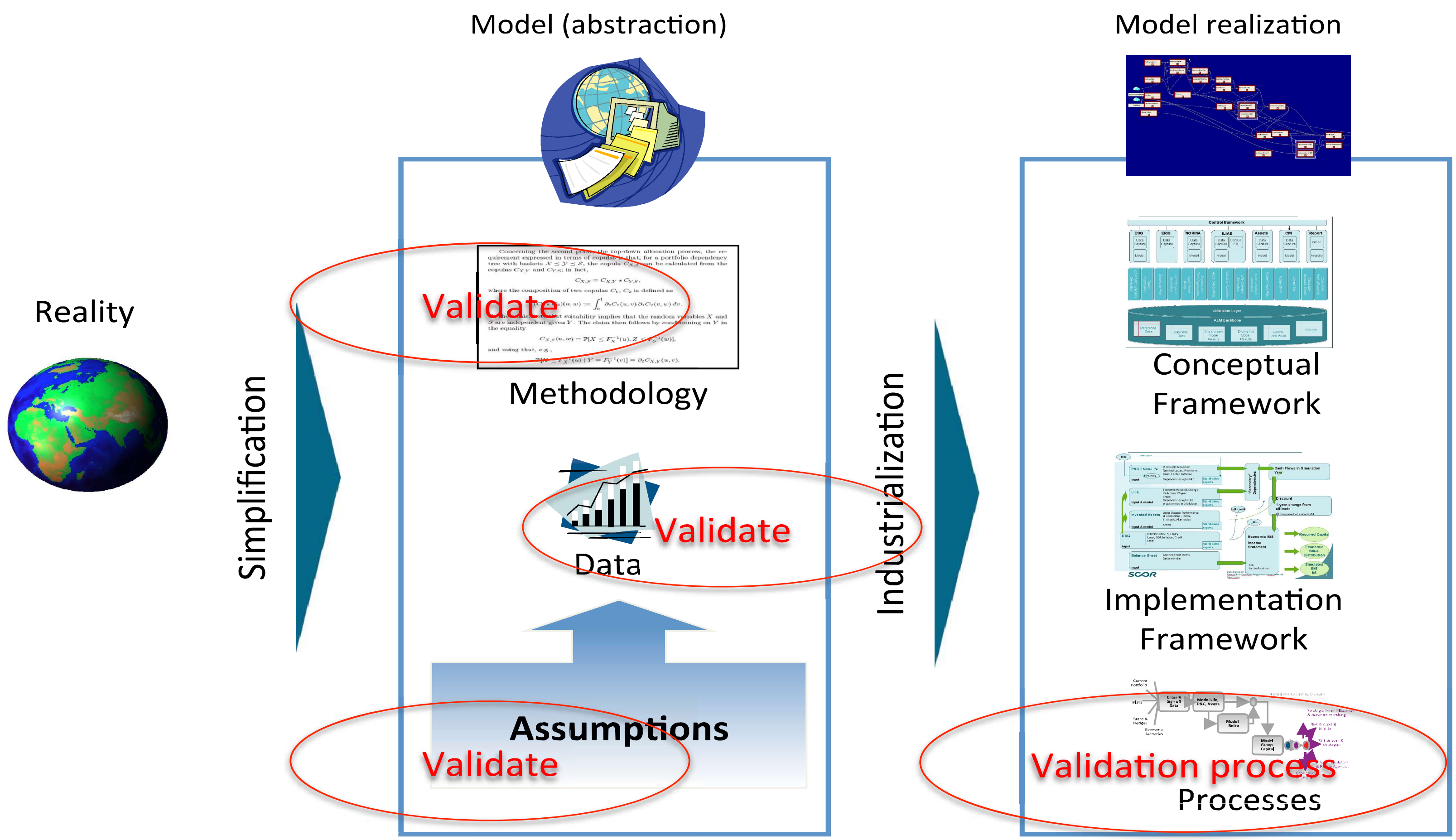
Validation of Risk Models

Marie KRATZ

ESSEC Business School
Paris Singapore

ERM 2016 - Conference of the Singapore Actuarial Society
Regent Hotel, Singapore, September 26-27, 2016

Internal models: validation



Model calibration

- Any model needs to determine few parameters. These parameters are set looking at data of the underlying process and fitting them to these data
- The pricing and reserving actuaries develop their model based on statistical tests on claims data
- The model is composed of probabilistic models for the various risk drivers but also to model for the dependence between those risks

  Both components need to be calibrated. The most difficult part being to find the right dependence between risks because this requires lots of data, particularly when there is only dependence in the tails (notion of non linear dependence)
- The probabilistic models are usually calibrated with claims data for the liabilities and with market data for the assets, or with simulations of stochastic models for events like natural catastrophes, pandemic, or for credit models

Model calibration

- Any model needs to determine few parameters. These parameters are set looking at data of the underlying process and fitting them to these data
- The pricing and reserving actuaries develop their model based on statistical tests on claims data
- The model is composed of probabilistic models for the various risk drivers but also to model for the dependence between those risks

  Both components need to be calibrated. The most difficult part being to find the right dependence between risks because this requires lots of data, particularly when there is only dependence in the tails (notion of non linear dependence)
- The probabilistic models are usually calibrated with claims data for the liabilities and with market data for the assets, or with simulations of stochastic models for events like natural catastrophes, pandemic, or for credit models

Model calibration

- Any model needs to determine few parameters. These parameters are set looking at data of the underlying process and fitting them to these data
- The pricing and reserving actuaries develop their model based on statistical tests on claims data
- The model is composed of probabilistic models for the various risk drivers but also to model for the dependence between those risks

  Both components need to be calibrated. The most difficult part being to find the right dependence between risks because this requires lots of data, particularly when there is only dependence in the tails (notion of non linear dependence)
- The probabilistic models are usually calibrated with claims data for the liabilities and with market data for the assets, or with simulations of stochastic models for events like natural catastrophes, pandemic, or for credit models

Model calibration

- Any model needs to determine few parameters. These parameters are set looking at data of the underlying process and fitting them to these data
- The pricing and reserving actuaries develop their model based on statistical tests on claims data
- The model is composed of probabilistic models for the various risk drivers but also to model for the dependence between those risks

  Both components need to be calibrated. The most difficult part being to find the right dependence between risks because this requires lots of data, particularly when there is only dependence in the tails (notion of non linear dependence)
- The probabilistic models are usually calibrated with claims data for the liabilities and with market data for the assets, or with simulations of stochastic models for events like natural catastrophes, pandemic, or for credit models

Model calibration

- Any model needs to determine few parameters. These parameters are set looking at data of the underlying process and fitting them to these data
- The pricing and reserving actuaries develop their model based on statistical tests on claims data
- The model is composed of probabilistic models for the various risk drivers but also to model for the dependence between those risks

  Both components need to be calibrated. The most difficult part being to find the right dependence between risks because this requires lots of data, particularly when there is only dependence in the tails (notion of non linear dependence)
- The probabilistic models are usually calibrated with claims data for the liabilities and with market data for the assets, or with simulations of stochastic models for events like natural catastrophes, pandemic, or for credit models

Model calibration : how to calibrate dependences ?

- Dependencies are NOT always linear : describing them by one number such as a linear correlation coefficient might be often oversimplifying !
- We generally use copulas to model dependences

Full knowledge of $(X_1, \cdots, X_n)$ =

knowledge of the joint distribution	knowledge of the margins
$F(x_1, \cdots, x_n) = \mathbb{P}[X_1 \leq x_1, \cdots, X_n \leq x_n]$	$F_i(x) = \mathbb{P}[X_1 \leq x], \forall i = 1, \cdots, n$
from which we deduce any marginal distribution :	+ knowledge of the
$F_i(x_i) = F(\infty, \cdots, \infty, x_i, \infty \cdots, \infty)$	dependence structure = copule

Sklar theorem

$$F(x_1, \cdots, x_n) = C(F_1(x_1), \cdots, F_n(x_n)), \quad \forall x_i \in \mathbb{R}, \quad i = 1, \cdots, n.$$

Model calibration : how to calibrate dependences ?

- Dependencies are NOT always linear : describing them by one number such as a linear correlation coefficient might be often oversimplifying !
- We generally use copulas to model dependences

Full knowledge of $(X_1, \cdots, X_n)$ =	
knowledge of the joint distribution $F(x_1, \cdots, x_n) = \mathbb{P}[X_1 \leq x_1, \cdots, X_n \leq x_n]$	knowledge of the margins $F_i(x) = \mathbb{P}[X_1 \leq x], \forall i = 1, \cdots, n$
from which we deduce any marginal distribution : $F_i(x_i) = F(\infty, \cdots, \infty, x_i, \infty \cdots, \infty)$	+ knowledge of the dependence structure = copule

Sklar theorem

$$F(x_1, \cdots, x_n) = C(F_1(x_1), \cdots, F_n(x_n)), \quad \forall x_i \in \mathbb{R}, \quad i = 1, \cdots, n.$$

Model calibration : various types of correlation

- Think using rank correlations (Kendall tau, Spearmans' rho) : Let C denote a copula of X and Y with parameter ρ.
  - Spearman's rho ρ_S : $\rho_S(X,Y) = \rho(F_X(X), F_Y(Y)) = \rho(\text{copula})$ and also $\rho_S(X,Y) = 12\int_0^1\int_0^1 \big(C(u,v) - u\,v\big)du\,dv.$
  - Kendall's tau ρ_τ : $\rho_\tau(X,Y) = 2\mathbb{P}\big[(X-\tilde{X})(Y-\tilde{Y}) > 0\big] - 1$ with $(\tilde{X},\tilde{Y})$ an independent copy of (X,Y), and also $\rho_\tau(X,Y) = 4\int_0^1\int_0^1 C(u,v)dC(u,v) - 1.$

Properties of Rank Correlation, NOT shared by Linear Correlation (enunciated for Spearman's rho ρ_S, but true also for Kendall's tau ρ_τ) :

1. ρ_S depends only on copula of $(X,Y)'$
2. ρ_S is invariant under strictly increasing transformations of the r.v.'s (contrarily to the case of linear correlation !)
3. $\rho_S(X,Y) = 1 \Leftrightarrow X, Y$ comonotonic
4. $\rho_S(X,Y) = -1 \Leftrightarrow X, Y$ countermonotonic.

Model calibration : various types of correlation

- Think using rank correlations (Kendall tau, Spearmans' rho) :
  Let C denote a copula of X and Y with parameter ρ.
  - Spearman's rho ρ_S : $\rho_S(X,Y) = \rho(F_X(X), F_Y(Y)) = \rho(\text{copula})$ and also $\rho_S(X,Y) = 12 \int_0^1 \int_0^1 \big(C(u,v) - u\,v\big) du\, dv.$
  - Kendall's tau ρ_τ : $\rho_\tau(X,Y) = 2\mathbb{P}\big[(X-\tilde{X})(Y-\tilde{Y}) > 0\big] - 1$ with $(\tilde{X},\tilde{Y})$ an independent copy of (X,Y), and also
    $\rho_\tau(X,Y) = 4 \int_0^1 \int_0^1 C(u,v) dC(u,v) - 1.$

Properties of Rank Correlation, NOT shared by Linear Correlation
(enunciated for Spearman's rho ρ_S, but true also for Kendall's tau ρ_τ) :

1. ρ_S depends only on copula of $(X,Y)'$
2. ρ_S is invariant under strictly increasing transformations of the r.v.'s (contrarily to the case of linear correlation !)
3. $\rho_S(X,Y) = 1 \Leftrightarrow X, Y$ comonotonic
4. $\rho_S(X,Y) = -1 \Leftrightarrow X, Y$ countermonotonic.

Model calibration : various types of correlation

- Think using rank correlations (Kendall tau, Spearmans' rho) : Let C denote a copula of X and Y with parameter ρ.
  - Spearman's rho ρ_S : $\rho_S(X,Y) = \rho(F_X(X), F_Y(Y)) = \rho(\text{copula})$ and also $\rho_S(X,Y) = 12\int_0^1\int_0^1 \big(C(u,v) - u\,v\big)du\,dv.$
  - Kendall's tau ρ_τ : $\rho_\tau(X,Y) = 2\mathbb{P}\big[(X-\tilde{X})(Y-\tilde{Y}) > 0\big] - 1$ with $(\tilde{X},\tilde{Y})$ an independent copy of (X,Y), and also

$$\rho_\tau(X,Y) = 4\int_0^1\int_0^1 C(u,v)dC(u,v) - 1.$$

Properties of Rank Correlation, NOT shared by Linear Correlation (enunciated for Spearman's rho ρ_S, but true also for Kendall's tau ρ_τ) :

1. ρ_S depends only on copula of $(X,Y)'$
2. ρ_S is invariant under strictly increasing transformations of the r.v.'s (contrarily to the case of linear correlation !)
3. $\rho_S(X,Y) = 1 \Leftrightarrow X, Y$ comonotonic
4. $\rho_S(X,Y) = -1 \Leftrightarrow X, Y$ countermonotonic.

Model calibration : various types of correlation

- Think using rank correlations (Kendall tau, Spearmans' rho) :
  Let C denote a copula of X and Y with parameter ρ.
  - Spearman's rho ρ_S : $\rho_S(X,Y) = \rho(F_X(X), F_Y(Y)) = \rho(\text{copula})$ and also $\rho_S(X,Y) = 12\int_0^1\int_0^1 \big(C(u,v) - u\,v\big)du\,dv.$
  - Kendall's tau ρ_τ : $\rho_\tau(X,Y) = 2\mathbb{P}\big[(X-\tilde{X})(Y-\tilde{Y}) > 0\big] - 1$ with $(\tilde{X},\tilde{Y})$ an independent copy of (X,Y), and also
    $\rho_\tau(X,Y) = 4\int_0^1\int_0^1 C(u,v)dC(u,v) - 1.$

Properties of Rank Correlation, NOT shared by Linear Correlation
(enunciated for Spearman's rho ρ_S, but true also for Kendall's tau ρ_τ) :

1. ρ_S depends only on copula of $(X,Y)'$
2. ρ_S is invariant under strictly increasing transformations of the r.v.'s (contrarily to the case of linear correlation !)
3. $\rho_S(X,Y) = 1 \Leftrightarrow X,Y$ comonotonic
4. $\rho_S(X,Y) = -1 \Leftrightarrow X,Y$ countermonotonic.

Model calibration : various types of correlation

- Think using rank correlations (Kendall tau, Spearmans' rho) : Let C denote a copula of X and Y with parameter ρ.
  - Spearman's rho ρ_S : $\rho_S(X,Y) = \rho(F_X(X), F_Y(Y)) = \rho(\text{copula})$ and also $\rho_S(X,Y) = 12\int_0^1\int_0^1 \big(C(u,v) - u\,v\big)du\,dv$.
  - Kendall's tau ρ_τ : $\rho_\tau(X,Y) = 2\mathbb{P}\big[(X-\tilde{X})(Y-\tilde{Y}) > 0\big] - 1$ with $(\tilde{X},\tilde{Y})$ an independent copy of (X,Y), and also $\rho_\tau(X,Y) = 4\int_0^1\int_0^1 C(u,v)dC(u,v) - 1$.

Properties of Rank Correlation, NOT shared by Linear Correlation
(enunciated for Spearman's rho ρ_S, but true also for Kendall's tau ρ_τ) :

1. ρ_S depends only on copula of $(X,Y)'$
2. ρ_S is invariant under strictly increasing transformations of the r.v.'s (contrarily to the case of linear correlation !)
3. $\rho_S(X,Y) = 1 \Leftrightarrow X, Y$ comonotonic
4. $\rho_S(X,Y) = -1 \Leftrightarrow X, Y$ countermonotonic.

Model calibration : tail dependence

- Think using coefficients of Tail dependence or Extremal dependence

  Objective : measure dependence in joint tail of bivariate distribution

  Coefficient of upper tail dependence : when limit exists, it is defined as

$$\lambda_u(X,Y) = \lim_{\alpha\to 1} \mathbb{P}\big[Y > VaR_\alpha(Y) \,|\, X > VaR_\alpha(X)\big]$$

and as function of the copula C of (X,Y),

$$\lambda_u(X,Y) = \lim_{\alpha\to 1} \frac{\overline{C}(\alpha,\alpha)}{1-\alpha} = \lim_{\alpha\to 1} \frac{1-2\alpha+C(\alpha,\alpha)}{1-\alpha}.$$

Coefficient of lower tail dependence : when limit exists, it is defined as $\lambda_l(X,Y) = \lim_{\alpha\to 0} \mathbb{P}\big[X \leq VaR_\alpha(X) \mid Y \leq VaR_\alpha(Y)\big]$ and as function of the copula, $\lambda_l(X,Y) = \lim_{\alpha\to 0} \frac{C(\alpha,\alpha)}{\alpha}$.

Model calibration : tail dependence

- Think using coefficients of Tail dependence or Extremal dependence

  Objective : measure dependence in joint tail of bivariate distribution

  Coefficient of upper tail dependence : when limit exists, it is defined as

$$\lambda_u(X,Y) = \lim_{\alpha\to 1} \mathbb{P}\big[Y > VaR_\alpha(Y) \,|\, X > VaR_\alpha(X)\big]$$

and as function of the copula C of (X,Y),

$$\lambda_u(X,Y) = \lim_{\alpha\to 1} \frac{\overline{C}(\alpha,\alpha)}{1-\alpha} = \lim_{\alpha\to 1} \frac{1-2\alpha+C(\alpha,\alpha)}{1-\alpha}.$$

Coefficient of lower tail dependence : when limit exists, it is defined as $\lambda_l(X,Y) = \lim_{\alpha\to 0} \mathbb{P}\big[X \le VaR_\alpha(X) \,|\, Y \le VaR_\alpha(Y)\big]$ and as function of the copula, $\lambda_l(X,Y) = \lim_{\alpha\to 0} \frac{C(\alpha,\alpha)}{\alpha}$.

Model calibration : tail dependence

- Think using coefficients of Tail dependence or Extremal dependence

  Objective : measure dependence in joint tail of bivariate distribution

  Coefficient of upper tail dependence : when limit exists, it is defined as

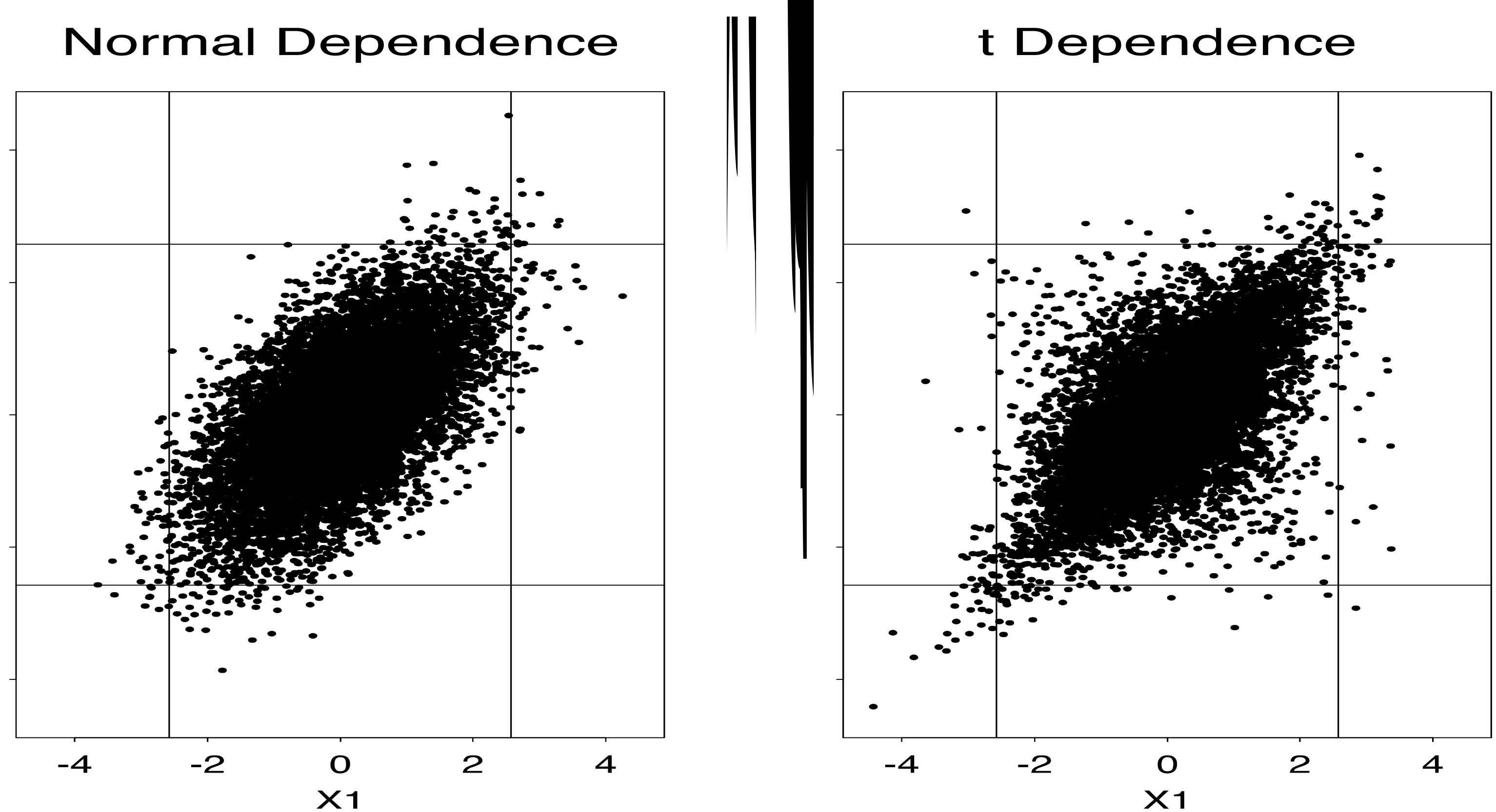
$$\lambda_u(X,Y) = \lim_{\alpha\to 1} \mathbb{P}\left[Y > VaR_\alpha(Y) \,|\, X > VaR_\alpha(X)\right]$$

and as function of the copula C of (X,Y),

$$\lambda_u(X,Y) = \lim_{\alpha\to 1} \frac{\overline{C}(\alpha,\alpha)}{1-\alpha} = \lim_{\alpha\to 1} \frac{1-2\alpha+C(\alpha,\alpha)}{1-\alpha}.$$

Coefficient of lower tail dependence : when limit exists, it is defined as $\lambda_l(X,Y) = \lim_{\alpha\to 0} \mathbb{P}\left[X \le VaR_\alpha(X) \mid Y \le VaR_\alpha(Y)\right]$ and as function of the copula, $\lambda_l(X,Y) = \lim_{\alpha\to 0} \frac{C(\alpha,\alpha)}{\alpha}$.

Gaussian and t3 Copulas Compared



Copula parameter $\rho = 0.7$; quantiles lines 0.5% and 99.5%.

Model calibration : how to calibrate dependences ?

- In insurance, there is often not enough liability data to estimate the copulas
- Nevertheless, copulas can be used to translate an expert opinion about dependencies in the portfolio into a model of dependence :
  - Select a copula with an appropriate shape : with increased dependences in the tail (this feature is observable in historic insurance loss data)
  - Try to estimate conditional probabilities by asking questions such as ’What about risk Y if risk X turned very bad ?’
    - Study the tail dependency (coefficients above)
    - Think about adverse scenarios in the portfolio
    - Look for causal relations between risks

Model calibration : how to calibrate dependences ?

- In insurance, there is often not enough liability data to estimate the copulas
- Nevertheless, copulas can be used to translate an expert opinion about dependencies in the portfolio into a model of dependence :
  - Select a copula with an appropriate shape : with increased dependences in the tail (this feature is observable in historic insurance loss data)
  - Try to estimate conditional probabilities by asking questions such as ’What about risk Y if risk X turned very bad ?’
    - Study the tail dependency (coefficients above)
    - Think about adverse scenarios in the portfolio
    - Look for causal relations between risks

Model calibration : how to calibrate dependences ?

- In insurance, there is often not enough liability data to estimate the copulas
- Nevertheless, copulas can be used to translate an expert opinion about dependencies in the portfolio into a model of dependence :
  - Select a copula with an appropriate shape : with increased dependences in the tail (this feature is observable in historic insurance loss data)
  - Try to estimate conditional probabilities by asking questions such as 'What about risk Y if risk X turned very bad ?'
    - Study the tail dependency (coefficients above)
    - Think about adverse scenarios in the portfolio
    - Look for causal relations between risks

Model calibration : how to calibrate dependences ?

- In insurance, there is often not enough liability data to estimate the copulas
- Nevertheless, copulas can be used to translate an expert opinion about dependencies in the portfolio into a model of dependence :
  - Select a copula with an appropriate shape : with increased dependences in the tail (this feature is observable in historic insurance loss data)
  - Try to estimate conditional probabilities by asking questions such as 'What about risk Y if risk X turned very bad ?'
    - Study the tail dependency (coefficients above)
    - Think about adverse scenarios in the portfolio
    - Look for causal relations between risks

Model calibration : Strategy for modeling dependences

- Using the knowledge of the underlying business to aggregate multiple risks, we can develop a hierarchical model for dependences in order to reduce the parameter space and describe more accurately the main sources of dependent behavior
- Once we have determined the structure of dependence for each node, there are two possibilities :
  - If we know a causal dependence, we model it explicitly
  - Otherwise, we systematically use non-symmetric copulas (ex. Clayton copula) in presence of tail dependence
- To calibrate the various nodes, we have again two possibilities :
  - If there is enough data, we calibrate statistically the parameters
  - In absence of data, we use stress scenarios or expert opinion to estimate conditional probabilities
- For the purpose of eliciting expert opinion (on common risk drivers, conditional probabilities, bucketing, $\cdots$), we can developed a Bayesian method combining various sources of information in the estimation. Example : PrObEx (see *Arbenz and Canestraro (2012). Estimating Copulas for Insurance from Scarce Observations, Expert Opinion and Prior Information : A Bayesian Approach. ASTIN Bulletin*)

Model calibration : Strategy for modeling dependences

- Using the knowledge of the underlying business to aggregate multiple risks, we can develop a hierarchical model for dependences in order to reduce the parameter space and describe more accurately the main sources of dependent behavior
- Once we have determined the structure of dependence for each node, there are two possibilities :
  - If we know a causal dependence, we model it explicitly
  - Otherwise, we systematically use non-symmetric copulas (ex. Clayton copula) in presence of tail dependence
- To calibrate the various nodes, we have again two possibilities :
  - If there is enough data, we calibrate statistically the parameters
  - In absence of data, we use stress scenarios or expert opinion to estimate conditional probabilities
- For the purpose of eliciting expert opinion (on common risk drivers, conditional probabilities, bucketing, $\cdots$), we can developed a Bayesian method combining various sources of information in the estimation. Example : PrObEx (see *Arbenz and Canestraro (2012). Estimating Copulas for Insurance from Scarce Observations, Expert Opinion and Prior Information : A Bayesian Approach. ASTIN Bulletin*)

Model calibration : Strategy for modeling dependences

- Using the knowledge of the underlying business to aggregate multiple risks, we can develop a hierarchical model for dependences in order to reduce the parameter space and describe more accurately the main sources of dependent behavior
- Once we have determined the structure of dependence for each node, there are two possibilities :
  - If we know a causal dependence, we model it explicitly
  - Otherwise, we systematically use non-symmetric copulas (ex. Clayton copula) in presence of tail dependence
- To calibrate the various nodes, we have again two possibilities :
  - If there is enough data, we calibrate statistically the parameters
  - In absence of data, we use stress scenarios or expert opinion to estimate conditional probabilities
- For the purpose of eliciting expert opinion (on common risk drivers, conditional probabilities, bucketing, $\cdots$), we can developed a Bayesian method combining various sources of information in the estimation. Example : PrObEx (see *Arbenz and Canestraro (2012). Estimating Copulas for Insurance from Scarce Observations, Expert Opinion and Prior Information : A Bayesian Approach. ASTIN Bulletin*)

Model calibration : Strategy for modeling dependences

- Using the knowledge of the underlying business to aggregate multiple risks, we can develop a hierarchical model for dependences in order to reduce the parameter space and describe more accurately the main sources of dependent behavior
- Once we have determined the structure of dependence for each node, there are two possibilities :
  - If we know a causal dependence, we model it explicitly
  - Otherwise, we systematically use non-symmetric copulas (ex. Clayton copula) in presence of tail dependence
- To calibrate the various nodes, we have again two possibilities :
  - If there is enough data, we calibrate statistically the parameters
  - In absence of data, we use stress scenarios or expert opinion to estimate conditional probabilities
- For the purpose of eliciting expert opinion (on common risk drivers, conditional probabilities, bucketing, $\cdots$), we can developed a Bayesian method combining various sources of information in the estimation. Example : PrObEx (see *Arbenz and Canestraro (2012). Estimating Copulas for Insurance from Scarce Observations, Expert Opinion and Prior Information : A Bayesian Approach. ASTIN Bulletin*)

Model calibration : Strategy for modeling dependences

- Using the knowledge of the underlying business to aggregate multiple risks, we can develop a hierarchical model for dependences in order to reduce the parameter space and describe more accurately the main sources of dependent behavior
- Once we have determined the structure of dependence for each node, there are two possibilities :
  - If we know a causal dependence, we model it explicitly
  - Otherwise, we systematically use non-symmetric copulas (ex. Clayton copula) in presence of tail dependence
- To calibrate the various nodes, we have again two possibilities :
  - If there is enough data, we calibrate statistically the parameters
  - In absence of data, we use stress scenarios or expert opinion to estimate conditional probabilities
- For the purpose of eliciting expert opinion (on common risk drivers, conditional probabilities, bucketing, $\cdots$), we can developed a Bayesian method combining various sources of information in the estimation. Example : PrObEx (see *Arbenz and Canestraro (2012). Estimating Copulas for Insurance from Scarce Observations, Expert Opinion and Prior Information : A Bayesian Approach. ASTIN Bulletin*)

Model calibration : Strategy for modeling dependences

- Using the knowledge of the underlying business to aggregate multiple risks, we can develop a hierarchical model for dependences in order to reduce the parameter space and describe more accurately the main sources of dependent behavior
- Once we have determined the structure of dependence for each node, there are two possibilities :
  - If we know a causal dependence, we model it explicitly
  - Otherwise, we systematically use non-symmetric copulas (ex. Clayton copula) in presence of tail dependence
- To calibrate the various nodes, we have again two possibilities :
  - If there is enough data, we calibrate statistically the parameters
  - In absence of data, we use stress scenarios or expert opinion to estimate conditional probabilities
- For the purpose of eliciting expert opinion (on common risk drivers, conditional probabilities, bucketing, $\cdots$), we can developed a Bayesian method combining various sources of information in the estimation. Example : PrObEx (see *Arbenz and Canestraro (2012). Estimating Copulas for Insurance from Scarce Observations, Expert Opinion and Prior Information : A Bayesian Approach. ASTIN Bulletin*)

Model calibration : Strategy for modeling dependences

- Using the knowledge of the underlying business to aggregate multiple risks, we can develop a hierarchical model for dependences in order to reduce the parameter space and describe more accurately the main sources of dependent behavior
- Once we have determined the structure of dependence for each node, there are two possibilities :
  - If we know a causal dependence, we model it explicitly
  - Otherwise, we systematically use non-symmetric copulas (ex. Clayton copula) in presence of tail dependence
- To calibrate the various nodes, we have again two possibilities :
  - If there is enough data, we calibrate statistically the parameters
  - In absence of data, we use stress scenarios or expert opinion to estimate conditional probabilities
- For the purpose of eliciting expert opinion (on common risk drivers, conditional probabilities, bucketing, $\cdots$), we can developed a Bayesian method combining various sources of information in the estimation. Example : PrObEx (see *Arbenz and Canestraro (2012). Estimating Copulas for Insurance from Scarce Observations, Expert Opinion and Prior Information : A Bayesian Approach. ASTIN Bulletin*)

Model calibration : Strategy for modeling dependences

- Using the knowledge of the underlying business to aggregate multiple risks, we can develop a hierarchical model for dependences in order to reduce the parameter space and describe more accurately the main sources of dependent behavior
- Once we have determined the structure of dependence for each node, there are two possibilities :
  - If we know a causal dependence, we model it explicitly
  - Otherwise, we systematically use non-symmetric copulas (ex. Clayton copula) in presence of tail dependence
- To calibrate the various nodes, we have again two possibilities :
  - If there is enough data, we calibrate statistically the parameters
  - In absence of data, we use stress scenarios or expert opinion to estimate conditional probabilities
- For the purpose of eliciting expert opinion (on common risk drivers, conditional probabilities, bucketing, $\cdots$), we can developed a Bayesian method combining various sources of information in the estimation. Example : PrObEx (see *Arbenz and Canestraro (2012). Estimating Copulas for Insurance from Scarce Observations, Expert Opinion and Prior Information : A Bayesian Approach. ASTIN Bulletin*)

Testing the various model components

Every internal model contains important components that will condition the results :

- An economic scenario generator / market risk
- A model for the uncertainty of P&C reserving triangles
- A model for natural catastrophes
- A model for pandemic (if there is a life book)
- A model for credit risk
- A model for operational risk
- A model for risk aggregation (dependence)

Each of these components can be tested independently, to check the validity of the methods employed

These tests vary from one component to the other. Each requires its own approach for testing

Testing the various model components

Every internal model contains important components that will condition the results :

- An economic scenario generator / market risk
- A model for the uncertainty of P&C reserving triangles
- A model for natural catastrophes
- A model for pandemic (if there is a life book)
- A model for credit risk
- A model for operational risk
- A model for risk aggregation (dependence)

Each of these components can be tested independently, to check the validity of the methods employed

These tests vary from one component to the other. Each requires its own approach for testing

Testing the various model components

Every internal model contains important components that will condition the results :

- An economic scenario generator / market risk
- A model for the uncertainty of P&C reserving triangles
- A model for natural catastrophes
- A model for pandemic (if there is a life book)
- A model for credit risk
- A model for operational risk
- A model for risk aggregation (dependence)

Each of these components can be tested independently, to check the validity of the methods employed

These tests vary from one component to the other. Each requires its own approach for testing

Example 1 : Testing the dependence model

▷ SCR depends crucially on the correct dependence model

- Using the wrong dependence model will lead to either an underestimation of the SCR (by neglecting the dependence in the tails) or an overestimation of the SCR (by fitting a correlation to a tail dependence as the Standard Formula does)
- Illustration : we compared statistics stemming from a 16-leaves full binary tree, when switching from lognormal(0,1) marginals and Flipped Clayton copulas with parameter $\theta = 1.36$, to Gaussian copulas calibrated either all in the extreme (same Quantile Exceedance Probability at 99,5% : 'tail correlation') or on the whole linear dependence (same 'Spearman correlation' coefficient 0.57)

Calibration	Capital ratio $VaR_{99.5}^{\text{Gauss}}/VaR_{99.5}^{\text{Clayton}}$
Gaussian copula with $\rho = 0.57$	0.64
Gaussian copula with $\rho = q^{Clayton}(0.05)$	1.06

Example 1 : Testing the dependence model

▷ SCR depends crucially on the correct dependence model

- Using the wrong dependence model will lead to either an underestimation of the SCR (by neglecting the dependence in the tails) or an overestimation of the SCR (by fitting a correlation to a tail dependence as the Standard Formula does)
- Illustration : we compared statistics stemming from a 16-leaves full binary tree, when switching from lognormal(0,1) marginals and Flipped Clayton copulas with parameter $\theta = 1.36$, to Gaussian copulas calibrated either all in the extreme (same Quantile Exceedance Probability at 99,5% : 'tail correlation') or on the whole linear dependence (same 'Spearman correlation' coefficient 0.57)

Calibration	Capital ratio $VaR_{99.5}^{\text{Gauss}}/VaR_{99.5}^{\text{Clayton}}$
Gaussian copula with $\rho = 0.57$	0.64
Gaussian copula with $\rho = q^{Clayton}(0.05)$	1.06

Example 2 : Testing the aggregation of the risk model
▷ Testing the Convergence of Monte Carlo Simulations

- We have developed a method to obtain explicit formulae for aggregated Pareto distributed risks linked by Clayton copula *(see M. Dacorogna, L. El Bahtouri, M. Kratz (2015-16). Explicit diversification benefit for dependent risks. SCOR Paper 38 & ESSEC WP 1522)*
- We use the results to test the convergence of the Monte Carlo simulations as a function of the parameters
- We compute both the TVaR (or ES) for the aggregated risks and the diversification benefit D of n dependent risks $(X_i)_{1\leq i\leq n}$ defined for a risk measure ρ as $D = 1 - \frac{\rho(\sum_{i=1}^{n} X_i)}{\sum_{i=1}^{n}\rho(X_i)}$
- We see that when the tail is very heavy, the simulations do not really converge

Example 2 : Testing the aggregation of the risk model
▷ Testing the Convergence of Monte Carlo Simulations

- We have developed a method to obtain explicit formulae for aggregated Pareto distributed risks linked by Clayton copula *(see M. Dacorogna, L. El Bahtouri, M. Kratz (2015-16). Explicit diversification benefit for dependent risks. SCOR Paper 38 & ESSEC WP 1522)*
- We use the results to test the convergence of the Monte Carlo simulations as a function of the parameters
- We compute both the TVaR (or ES) for the aggregated risks and the diversification benefit D of n dependent risks $(X_i)_{1\leq i\leq n}$ defined for a risk measure ρ as $D = 1 - \frac{\rho(\sum_{i=1}^n X_i)}{\sum_{i=1}^n \rho(X_i)}$
- We see that when the tail is very heavy, the simulations do not really converge

Example 2 : Testing the aggregation of the risk model
▷ Testing the Convergence of Monte Carlo Simulations

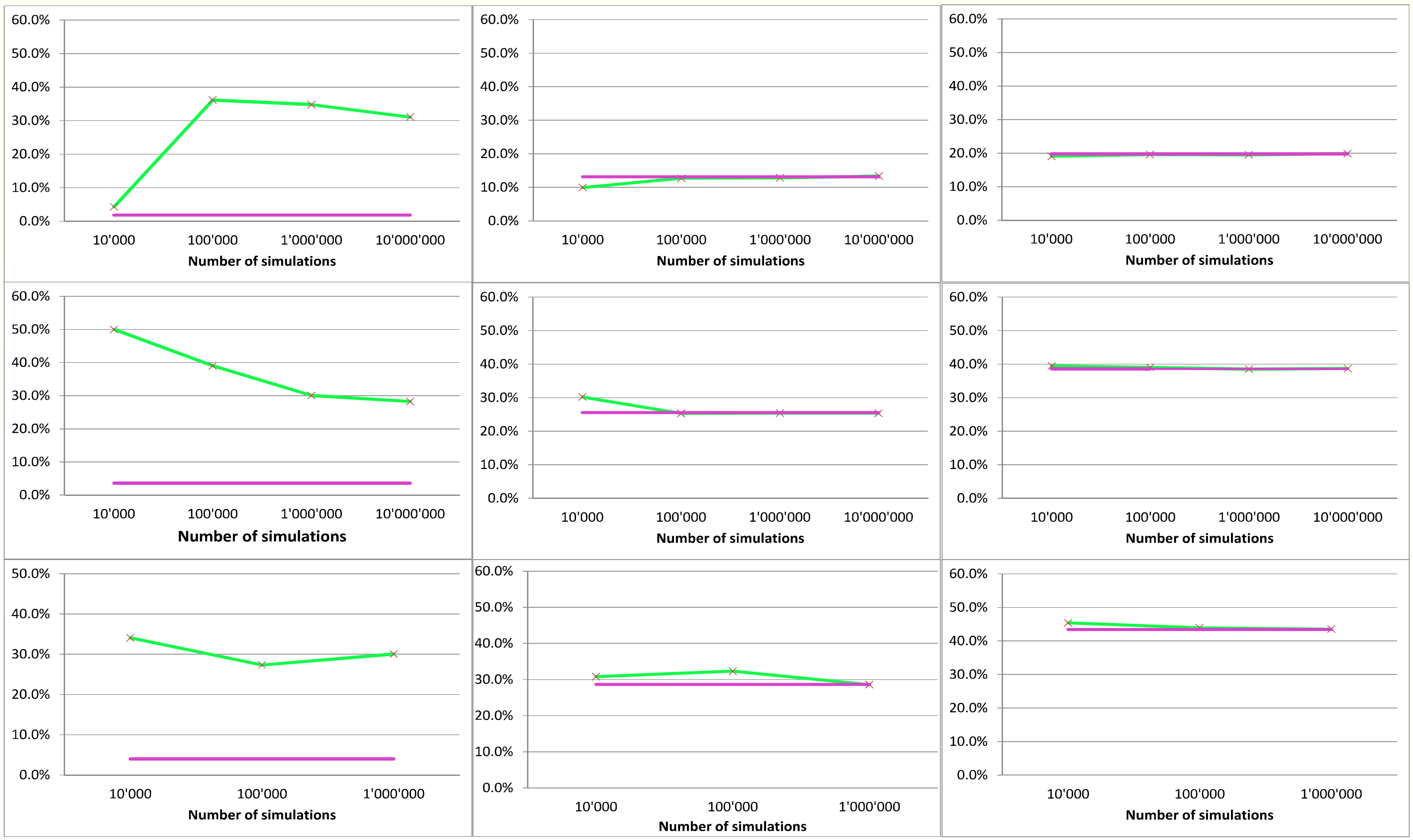
- We have developed a method to obtain explicit formulae for aggregated Pareto distributed risks linked by Clayton copula *(see M. Dacorogna, L. El Bahtouri, M. Kratz (2015-16). Explicit diversification benefit for dependent risks. SCOR Paper 38 & ESSEC WP 1522)*
- We use the results to test the convergence of the Monte Carlo simulations as a function of the parameters
- We compute both the TVaR (or ES) for the aggregated risks and the diversification benefit D of n dependent risks $(X_i)_{1\leq i\leq n}$ defined for a risk measure ρ as $D = 1 - \frac{\rho(\sum_{i=1}^n X_i)}{\sum_{i=1}^n \rho(X_i)}$
- We see that when the tail is very heavy, the simulations do not really converge

Example 2 : Testing the aggregation of the risk model
▷ Testing the Convergence of Monte Carlo Simulations

- We have developed a method to obtain explicit formulae for aggregated Pareto distributed risks linked by Clayton copula *(see M. Dacorogna, L. El Bahtouri, M. Kratz (2015-16). Explicit diversification benefit for dependent risks. SCOR Paper 38 & ESSEC WP 1522)*
- We use the results to test the convergence of the Monte Carlo simulations as a function of the parameters
- We compute both the TVaR (or ES) for the aggregated risks and the diversification benefit D of n dependent risks $(X_i)_{1\le i\le n}$ defined for a risk measure ρ as $D = 1 - \dfrac{\rho(\sum_{i=1}^{n} X_i)}{\sum_{i=1}^{n} \rho(X_i)}$
- We see that when the tail is very heavy, the simulations do not really converge

Convergence of Diversification Benefit : Pareto (α) risks with Survival Clayton (θ) Copula. From up to bottom : n = 2, 10, 100 respectively.

$\alpha = 1.1\,;\ \theta = 0.91$ $\quad\quad$ $\alpha = 2\,;\ \theta = 0.5$ $\quad\quad$ $\alpha = 3\,;\ \theta = 0.3$



Good convergence for $\alpha = 2$ or 3, but it does not converge for $\alpha = 1.1$.

Example 3 : Testing the market risk model : Popular backtesting for risk measures

▷ for VaR : Binomial test on the proportion p of VaR violations,

estimated by $\frac{1}{n}\sum_{t=1}^{n} I_t(\alpha)$, with $I_t(\alpha) = \mathbf{1}_{\{L(t) > VaR_\alpha(L(t))\}}$:

$$H0 : p = p_0 := 1-\alpha \quad \text{against} \quad H1 : p > p_0$$

If the proportion of VaR violations is not significantly different from $1-\alpha$, then the estimation/prediction method is reasonable.

▷ for TVaR (or ES) (*see M.Kratz, Y. Lok, A. McNeil (2016). An implicit backtest for ES via a simple multinomial approach. ArXiv* : Main idea= backtesting ES via simultaneously backtesting multiple VaR estimates

$$\begin{cases} H0 : & \beta_j = \alpha_j \quad \text{for } j = 1,\ldots,N \\ H1 : & \beta_j \neq \alpha_j \quad \text{for at least one } j \in \{1,\ldots,N\}. \end{cases}$$

with $O_j = \sum_{t=1}^{n} I_{(q_{j-1} < L_t \leq q_j)}$, for $j = 1, ..., N+1$,

$(O_1,\ldots,O_{N+1}) \sim \text{MN}(\beta_1-\beta_0,\ldots,\beta_{N+1}-\beta_N)$ for parameters $\beta_1 < \cdots < \beta_N$ with $\beta_0 = 0$ and $\beta_{N+1} = 1$.

Example 3 : Testing the market risk model : Popular backtesting for risk measures

▷ for VaR : Binomial test on the proportion p of VaR violations, estimated by $\frac{1}{n}\sum_{t=1}^{n} I_t(\alpha)$, with $I_t(\alpha) = \mathbf{1}_{\{L(t) > VaR_\alpha(L(t))\}}$:

H0 : $p = p_0 := 1 - \alpha$ against H1 : $p > p_0$

If the proportion of VaR violations is not significantly different from $1-\alpha$, then the estimation/prediction method is reasonable.

▷ for TVaR (or ES) (*see M.Kratz, Y. Lok, A. McNeil (2016). An implicit backtest for ES via a simple multinomial approach. ArXiv* : Main idea= backtesting ES via simultaneously backtesting multiple VaR estimates

$$\left|\begin{array}{llll} \text{H0}: & \beta_j & = \alpha_j & \text{for } j = 1,\dots,N \\ \text{H1}: & \beta_j & \neq \alpha_j & \text{for at least one } j \in \{1,\dots,N\}. \end{array}\right.$$

with $O_j = \sum_{t=1}^{n} I_{(q_{j-1} < L_t \le q_j)}$, for $j = 1, ..., N+1$,

$(O_1, \dots, O_{N+1}) \sim \text{MN}(\beta_1 - \beta_0, \dots, \beta_{N+1} - \beta_N)$ for parameters $\beta_1 < \cdots < \beta_N$ with $\beta_0 = 0$ and $\beta_{N+1} = 1$.

Example

TABLE: Rejection rate for the null hypothesis (H0) on a sample size of length n_1, using a multinomial approach with 3 possible tests (χ^2, Nass, LR) to backtest simultaneously the $N = 2^k$, $1 \leq k \leq 6$, quantiles VaR_{α_j}, $1 \leq j \leq N$, with $\alpha_1 = \alpha = 97.5\%$, on data simulated from various distributions (normal, Student $t3$, $t5$ and skewed $t3$)

		Chi Square							Nass							LR						
	n1	1	2	4	8	16	32	64	1	2	4	8	16	32	64	1	2	4	8	16	32	64
Standard Normal	250	4,1%	5,0%	5,3%	9,1%	11,3%	15,0%	22,3%	4,1%	3,7%	5,0%	5,3%	5,6%	5,3%	5,2%	7,7%	10,4%	6,3%	6,2%	6,4%	6,2%	6,0%
	500	4,4%	4,8%	5,2%	6,2%	8,4%	11,8%	15,7%	4,4%	4,2%	4,6%	4,6%	5,4%	5,4%	5,0%	6,5%	5,9%	5,7%	5,7%	5,5%	5,5%	5,4%
	1000	5,1%	4,6%	5,2%	5,9%	7,4%	9,2%	12,3%	5,1%	4,2%	4,9%	5,1%	5,3%	5,4%	5,4%	4,2%	5,3%	5,5%	5,4%	5,2%	5,2%	5,2%
	2000	5,2%	4,9%	5,0%	5,7%	6,3%	7,4%	9,7%	5,2%	4,8%	4,6%	5,1%	5,4%	5,3%	5,6%	4,3%	5,3%	5,0%	4,9%	4,9%	4,9%	4,9%
t5	250	5,0%	10,6%	14,5%	21,9%	23,1%	27,4%	34,0%	5,0%	8,3%	13,2%	14,8%	14,3%	14,8%	13,7%	8,0%	15,6%	16,6%	22,6%	27,0%	31,1%	34,1%
	500	5,4%	15,8%	22,1%	28,5%	31,8%	36,1%	39,0%	5,4%	14,5%	20,1%	24,4%	26,4%	25,4%	22,7%	6,8%	16,0%	26,6%	36,9%	44,7%	50,3%	54,5%
	1000	6,6%	27,5%	41,7%	49,8%	54,1%	55,0%	55,6%	6,6%	26,4%	40,9%	47,2%	49,9%	48,0%	43,7%	5,0%	26,9%	48,3%	63,0%	72,4%	78,3%	81,4%
	2000	7,5%	47,9%	71,0%	79,7%	82,7%	82,8%	81,6%	7,5%	47,8%	70,2%	78,7%	81,2%	79,8%	76,7%	6,0%	48,9%	77,4%	89,7%	94,5%	96,7%	97,8%
t3	250	3,8%	7,1%	13,3%	20,8%	19,5%	25,6%	28,3%	3,8%	5,3%	11,7%	14,2%	13,8%	13,8%	13,9%	10,3%	24,7%	24,3%	35,6%	42,2%	48,1%	52,1%
	500	5,3%	16,0%	24,3%	32,5%	34,4%	39,6%	38,5%	5,3%	15,4%	21,4%	27,8%	31,6%	28,9%	25,8%	9,8%	27,1%	44,7%	58,8%	68,1%	73,9%	77,7%
	1000	9,9%	37,7%	56,5%	63,6%	65,4%	64,4%	64,4%	9,9%	35,6%	55,2%	60,9%	62,0%	60,0%	54,3%	9,8%	47,6%	75,3%	88,0%	93,0%	95,6%	96,6%
	2000	17,4%	73,7%	90,9%	94,6%	94,8%	93,7%	91,9%	17,4%	73,3%	90,4%	94,0%	94,2%	92,5%	89,6%	17,4%	80,3%	96,7%	99,3%	99,8%	100,0%	100,0%
skt3	250	13,6%	38,7%	52,0%	64,0%	63,5%	69,8%	73,8%	13,6%	34,2%	49,5%	54,6%	55,0%	55,2%	55,0%	14,3%	34,8%	53,5%	66,3%	73,9%	78,4%	81,6%
	500	24,0%	63,3%	79,0%	85,7%	88,1%	89,8%	90,7%	24,0%	60,8%	77,2%	82,9%	86,0%	85,2%	84,1%	24,1%	58,6%	81,7%	90,7%	94,4%	96,4%	97,2%
	1000	41,7%	89,4%	97,1%	98,7%	99,2%	99,3%	99,3%	41,7%	89,0%	96,9%	98,6%	99,1%	99,0%	98,8%	35,2%	87,5%	97,9%	99,6%	99,8%	100,0%	100,0%
	2000	66,2%	99,6%	100,0%	100,0%	100,0%	100,0%	100,0%	66,2%	99,6%	100,0%	100,0%	100,0%	100,0%	100,0%	61,6%	99,5%	100,0%	100,0%	100,0%	100,0%	100,0%

Is it possible to statistically test internal models ?

- RAC (Risk Adjusted Capital) is computed for a probability of 1% or 0.5%, which represents a 1/100 or 1/200 years event
- In most of the insured risks, such an event has never been observed or has been observed only once
- This means that the tails of the distributions have to be inferred from data from the last 10 to 30 years in the best cases
- The 1/100 years RAC is thus based on a theoretical estimate of the shock (event) size
- It is considered more as the rule of the game than as a realistic risk cover
- It is a compromise between pure betting and not doing anything because we cannot statistically estimate it

Is it possible to statistically test internal models ?

- RAC (Risk Adjusted Capital) is computed for a probability of 1% or 0.5%, which represents a 1/100 or 1/200 years event
- In most of the insured risks, such an event has never been observed or has been observed only once
- This means that the tails of the distributions have to be inferred from data from the last 10 to 30 years in the best cases
- The 1/100 years RAC is thus based on a theoretical estimate of the shock (event) size
- It is considered more as the rule of the game than as a realistic risk cover
- It is a compromise between pure betting and not doing anything because we cannot statistically estimate it

Is it possible to statistically test internal models ?

- RAC (Risk Adjusted Capital) is computed for a probability of 1% or 0.5%, which represents a 1/100 or 1/200 years event
- In most of the insured risks, such an event has never been observed or has been observed only once
- This means that the tails of the distributions have to be inferred from data from the last 10 to 30 years in the best cases
- The 1/100 years RAC is thus based on a theoretical estimate of the shock (event) size
- It is considered more as the rule of the game than as a realistic risk cover
- It is a compromise between pure betting and not doing anything because we cannot statistically estimate it

Is it possible to statistically test internal models ?

- RAC (Risk Adjusted Capital) is computed for a probability of 1% or 0.5%, which represents a 1/100 or 1/200 years event
- In most of the insured risks, such an event has never been observed or has been observed only once
- This means that the tails of the distributions have to be inferred from data from the last 10 to 30 years in the best cases
- The 1/100 years RAC is thus based on a theoretical estimate of the shock (event) size
- It is considered more as the rule of the game than as a realistic risk cover
- It is a compromise between pure betting and not doing anything because we cannot statistically estimate it

Is it possible to statistically test internal models ?

- RAC (Risk Adjusted Capital) is computed for a probability of 1% or 0.5%, which represents a 1/100 or 1/200 years event
- In most of the insured risks, such an event has never been observed or has been observed only once
- This means that the tails of the distributions have to be inferred from data from the last 10 to 30 years in the best cases
- The 1/100 years RAC is thus based on a theoretical estimate of the shock (event) size
- It is considered more as the rule of the game than as a realistic risk cover
- It is a compromise between pure betting and not doing anything because we cannot statistically estimate it

Is it possible to statistically test internal models ?

- RAC (Risk Adjusted Capital) is computed for a probability of 1% or 0.5%, which represents a 1/100 or 1/200 years event
- In most of the insured risks, such an event has never been observed or has been observed only once
- This means that the tails of the distributions have to be inferred from data from the last 10 to 30 years in the best cases
- The 1/100 years RAC is thus based on a theoretical estimate of the shock (event) size
- It is considered more as the rule of the game than as a realistic risk cover
- It is a compromise between pure betting and not doing anything because we cannot statistically estimate it

A way to check the validity of the model : Stress testing

- Testing the output of internal models is thus a must to gain confidence in its results and to understand its limitations
- We just saw that it is difficult, or even impossible, to statistically test the model. We can only stress test it
- There are at least four ways of stress testing the models :
  1. Test the sensitivity to parameters (sensitivity analysis)
  2. Test the predictions against real outcomes (historical test, via P&L attribution for lines of business (LoB) and assets)
  3. Test the model against scenarios that can be seen as thought experiments about possible future world situations. By comparing the probability of the scenario given by the internal model to the expected frequency of such a scenario, we can assess whether the internal model is realistic and has really taken into account enough dependencies between risks
  4. Study the reasonableness of the extreme scenarios of the Monte-Carlo simulations (reverse stress-test)

A way to check the validity of the model : Stress testing

- Testing the output of internal models is thus a must to gain confidence in its results and to understand its limitations
- We just saw that it is difficult, or even impossible, to statistically test the model. We can only stress test it
- There are at least four ways of stress testing the models :
  1. Test the sensitivity to parameters (sensitivity analysis)
  2. Test the predictions against real outcomes (historical test, via P&L attribution for lines of business (LoB) and assets)
  3. Test the model against scenarios that can be seen as thought experiments about possible future world situations. By comparing the probability of the scenario given by the internal model to the expected frequency of such a scenario, we can assess whether the internal model is realistic and has really taken into account enough dependencies between risks
  4. Study the reasonableness of the extreme scenarios of the Monte-Carlo simulations (reverse stress-test)

A way to check the validity of the model : Stress testing

- Testing the output of internal models is thus a must to gain confidence in its results and to understand its limitations
- We just saw that it is difficult, or even impossible, to statistically test the model. We can only stress test it
- There are at least four ways of stress testing the models :
  1. Test the sensitivity to parameters (sensitivity analysis)
  2. Test the predictions against real outcomes (historical test, via P&L attribution for lines of business (LoB) and assets)
  3. Test the model against scenarios that can be seen as thought experiments about possible future world situations. By comparing the probability of the scenario given by the internal model to the expected frequency of such a scenario, we can assess whether the internal model is realistic and has really taken into account enough dependencies between risks
  4. Study the reasonableness of the extreme scenarios of the Monte-Carlo simulations (reverse stress-test)

A way to check the validity of the model : Stress testing

- Testing the output of internal models is thus a must to gain confidence in its results and to understand its limitations
- We just saw that it is difficult, or even impossible, to statistically test the model. We can only stress test it
- There are at least four ways of stress testing the models :
  1. Test the sensitivity to parameters (sensitivity analysis)
  2. Test the predictions against real outcomes (historical test, via P&L attribution for lines of business (LoB) and assets)
  3. Test the model against scenarios that can be seen as thought experiments about possible future world situations. By comparing the probability of the scenario given by the internal model to the expected frequency of such a scenario, we can assess whether the internal model is realistic and has really taken into account enough dependencies between risks
  4. Study the reasonableness of the extreme scenarios of the Monte-Carlo simulations (reverse stress-test)

A way to check the validity of the model : Stress testing

- Testing the output of internal models is thus a must to gain confidence in its results and to understand its limitations
- We just saw that it is difficult, or even impossible, to statistically test the model. We can only stress test it
- There are at least four ways of stress testing the models :
  1. Test the sensitivity to parameters (sensitivity analysis)
  2. Test the predictions against real outcomes (historical test, via P&L attribution for lines of business (LoB) and assets)
  3. Test the model against scenarios that can be seen as thought experiments about possible future world situations. By comparing the probability of the scenario given by the internal model to the expected frequency of such a scenario, we can assess whether the internal model is realistic and has really taken into account enough dependencies between risks
  4. Study the reasonableness of the extreme scenarios of the Monte-Carlo simulations (reverse stress-test)

A way to check the validity of the model : Stress testing

- Testing the output of internal models is thus a must to gain confidence in its results and to understand its limitations
- We just saw that it is difficult, or even impossible, to statistically test the model. We can only stress test it
- There are at least four ways of stress testing the models :
  1. Test the sensitivity to parameters (sensitivity analysis)
  2. Test the predictions against real outcomes (historical test, via P&L attribution for lines of business (LoB) and assets)
  3. Test the model against scenarios that can be seen as thought experiments about possible future world situations. By comparing the probability of the scenario given by the internal model to the expected frequency of such a scenario, we can assess whether the internal model is realistic and has really taken into account enough dependencies between risks
  4. Study the reasonableness of the extreme scenarios of the Monte-Carlo simulations (reverse stress-test)

A way to check the validity of the model : Stress testing

- Testing the output of internal models is thus a must to gain confidence in its results and to understand its limitations
- We just saw that it is difficult, or even impossible, to statistically test the model. We can only stress test it
- There are at least four ways of stress testing the models :
  1. Test the sensitivity to parameters (sensitivity analysis)
  2. Test the predictions against real outcomes (historical test, via P&L attribution for lines of business (LoB) and assets)
  3. Test the model against scenarios that can be seen as thought experiments about possible future world situations. By comparing the probability of the scenario given by the internal model to the expected frequency of such a scenario, we can assess whether the internal model is realistic and has really taken into account enough dependencies between risks
  4. Study the reasonableness of the extreme scenarios of the Monte-Carlo simulations (reverse stress-test)

Another way to look at the quality of the model : Reverse stress test

Making full use of the Monte Carlo simulations

- Stochastic models produce many simulations at each run. These outputs can be put at use to understand the way the model works
- We select the worst cases and look at what are the scenarios that make the company bankrupted.
- Two questions to ask on these scenarios :
  - Is this scenario credible given the company portfolio ?
  - Are there other possible scenarios that do not appear in the worst Monte Carlo simulations ?
- This is typically the kind of reverse back testing that can be done on the simulations
- Other tests are also interesting like looking at conditional statistics. A typical question would for instance be : how is the capital going to behave if interest rises ?

Another way to look at the quality of the model : Reverse stress test
Making full use of the Monte Carlo simulations

- Stochastic models produce many simulations at each run. These outputs can be put at use to understand the way the model works
- We select the worst cases and look at what are the scenarios that make the company bankrupted.
- Two questions to ask on these scenarios :
  - Is this scenario credible given the company portfolio ?
  - Are there other possible scenarios that do not appear in the worst Monte Carlo simulations ?
- This is typically the kind of reverse back testing that can be done on the simulations
- Other tests are also interesting like looking at conditional statistics. A typical question would for instance be : how is the capital going to behave if interest rises ?

Another way to look at the quality of the model : Reverse stress test

Making full use of the Monte Carlo simulations

- Stochastic models produce many simulations at each run. These outputs can be put at use to understand the way the model works
- We select the worst cases and look at what are the scenarios that make the company bankrupted.
- Two questions to ask on these scenarios :
  - Is this scenario credible given the company portfolio ?
  - Are there other possible scenarios that do not appear in the worst Monte Carlo simulations ?
- This is typically the kind of reverse back testing that can be done on the simulations
- Other tests are also interesting like looking at conditional statistics. A typical question would for instance be : how is the capital going to behave if interest rises ?

Another way to look at the quality of the model : Reverse stress test
Making full use of the Monte Carlo simulations

- Stochastic models produce many simulations at each run. These outputs can be put at use to understand the way the model works
- We select the worst cases and look at what are the scenarios that make the company bankrupted.
- Two questions to ask on these scenarios :
  - Is this scenario credible given the company portfolio ?
  - Are there other possible scenarios that do not appear in the worst Monte Carlo simulations ?
- This is typically the kind of reverse back testing that can be done on the simulations
- Other tests are also interesting like looking at conditional statistics. A typical question would for instance be : how is the capital going to behave if interest rises ?

Another way to look at the quality of the model : Reverse stress test
Making full use of the Monte Carlo simulations

- Stochastic models produce many simulations at each run. These outputs can be put at use to understand the way the model works
- We select the worst cases and look at what are the scenarios that make the company bankrupted.
- Two questions to ask on these scenarios :
  - Is this scenario credible given the company portfolio ?
  - Are there other possible scenarios that do not appear in the worst Monte Carlo simulations ?
- This is typically the kind of reverse back testing that can be done on the simulations
- Other tests are also interesting like looking at conditional statistics. A typical question would for instance be : how is the capital going to behave if interest rises ?

Another way to look at the quality of the model : Reverse stress test

Making full use of the Monte Carlo simulations

- Stochastic models produce many simulations at each run. These outputs can be put at use to understand the way the model works
- We select the worst cases and look at what are the scenarios that make the company bankrupted.
- Two questions to ask on these scenarios :
  - Is this scenario credible given the company portfolio ?
  - Are there other possible scenarios that do not appear in the worst Monte Carlo simulations ?
- This is typically the kind of reverse back testing that can be done on the simulations
- Other tests are also interesting like looking at conditional statistics. A typical question would for instance be : how is the capital going to behave if interest rises ?

Another way to look at the quality of the model : Reverse stress test
Making full use of the Monte Carlo simulations

- Stochastic models produce many simulations at each run. These outputs can be put at use to understand the way the model works
- We select the worst cases and look at what are the scenarios that make the company bankrupted.
- Two questions to ask on these scenarios :
  - Is this scenario credible given the company portfolio ?
  - Are there other possible scenarios that do not appear in the worst Monte Carlo simulations ?
- This is typically the kind of reverse back testing that can be done on the simulations
- Other tests are also interesting like looking at conditional statistics. A typical question would for instance be : how is the capital going to behave if interest rises ?

Example of Reverse stress-test

- Internal models generate a huge quantity of data. Usually little of these data is used : some averages for computing capital and some expectations
- Exploring the dependence of results to certain important variables is a very good way to test the reasonableness of the model
- In the next few slides, we present regression plots, which show the dependency between interest rates and change in economic value (of certain LoB's)
- The plots are based on the full 100'000 scenarios of an Internal Model
- By analyzing the model results on this level, we can follow up on a lot of effects and test if they make sense

Example of Reverse stress-test

- Internal models generate a huge quantity of data. Usually little of these data is used : some averages for computing capital and some expectations
- Exploring the dependence of results to certain important variables is a very good way to test the reasonableness of the model
- In the next few slides, we present regression plots, which show the dependency between interest rates and change in economic value (of certain LoB's)
- The plots are based on the full 100'000 scenarios of an Internal Model
- By analyzing the model results on this level, we can follow up on a lot of effects and test if they make sense

Example of Reverse stress-test

- Internal models generate a huge quantity of data. Usually little of these data is used : some averages for computing capital and some expectations
- Exploring the dependence of results to certain important variables is a very good way to test the reasonableness of the model
- In the next few slides, we present regression plots, which show the dependency between interest rates and change in economic value (of certain LoB's)
- The plots are based on the full 100'000 scenarios of an Internal Model
- By analyzing the model results on this level, we can follow up on a lot of effects and test if they make sense

Example of Reverse stress-test

- Internal models generate a huge quantity of data. Usually little of these data is used : some averages for computing capital and some expectations
- Exploring the dependence of results to certain important variables is a very good way to test the reasonableness of the model
- In the next few slides, we present regression plots, which show the dependency between interest rates and change in economic value (of certain LoB's)
- The plots are based on the full 100'000 scenarios of an Internal Model
- By analyzing the model results on this level, we can follow up on a lot of effects and test if they make sense

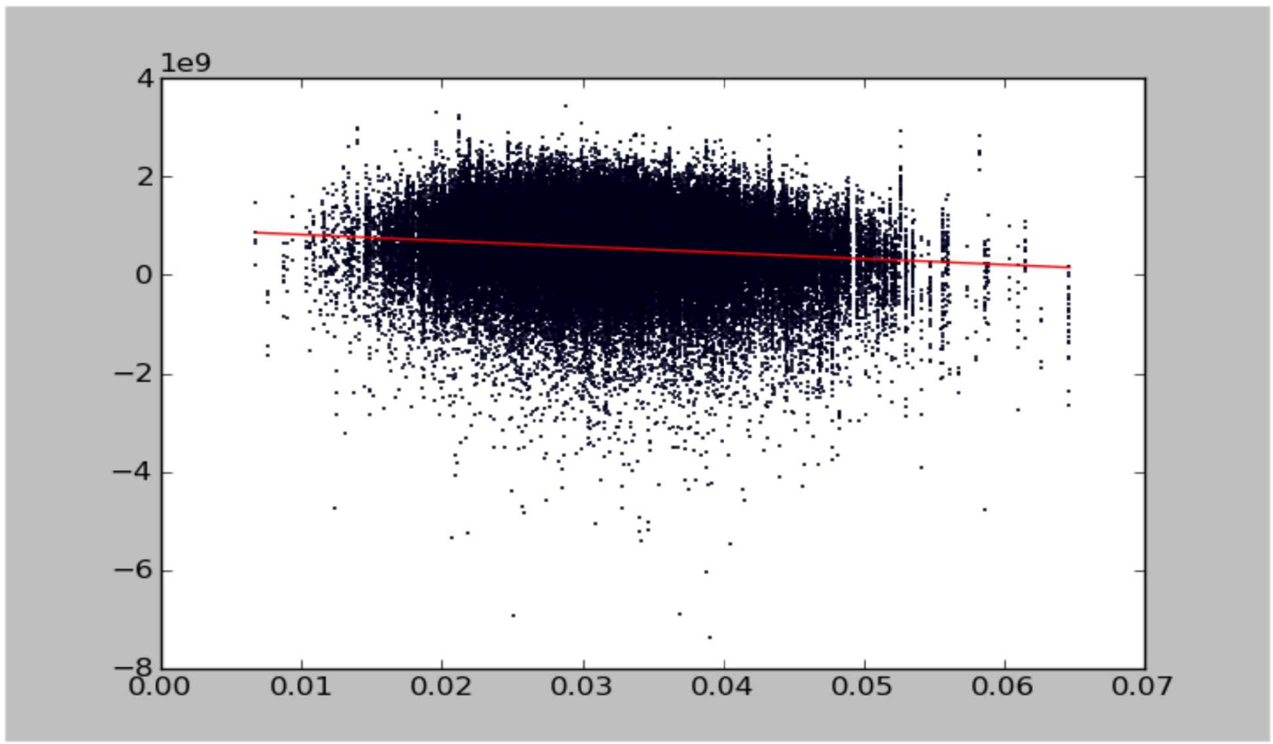
Example of Reverse stress-test

- Internal models generate a huge quantity of data. Usually little of these data is used : some averages for computing capital and some expectations
- Exploring the dependence of results to certain important variables is a very good way to test the reasonableness of the model
- In the next few slides, we present regression plots, which show the dependency between interest rates and change in economic value (of certain LoB's)
- The plots are based on the full 100'000 scenarios of an Internal Model
- By analyzing the model results on this level, we can follow up on a lot of effects and test if they make sense

Example :

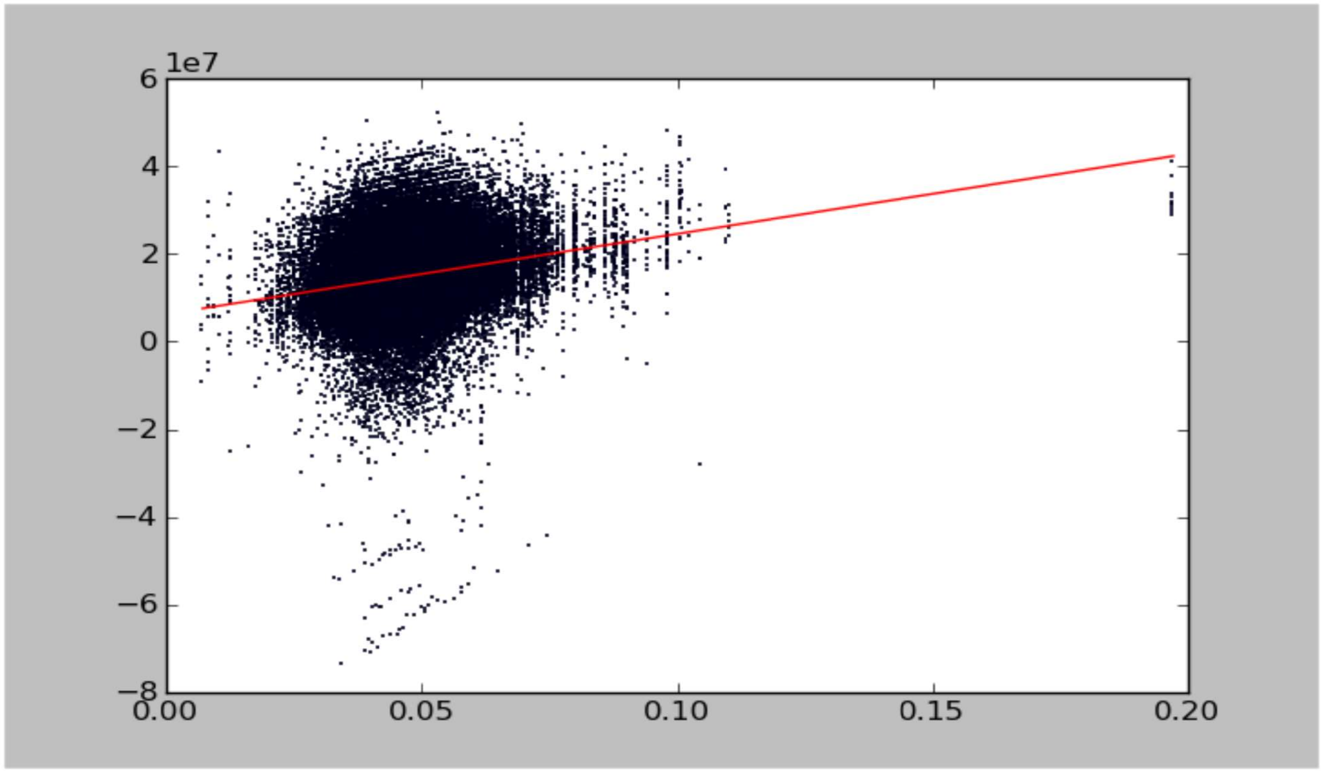
Change of Company Value versus the 4Y EUR Gov.

4Y is the typical duration of the P&C portfolio



As interest rate grows the Value of the company slightly decreases (due to an increase in inflation linked to IR increase)

Professional Liability (long tail) versus 5Y GBP



The value of professional liability business depends heavily on interest rate as it takes a long time to develop to ultimate and the reserve can earn interest for a longer time

Conclusion

- The development of risk models helps to improve risk awareness and anchors risk management and governance deeper in industry practices
- Risk models provide valuable assessments, especially in relative terms, as well as guidance in business decisions
- It is thus essential to ensure that the results of the model delivers a good description of reality
- Model validation is the way to gain confidence in the model
- It is however difficult because there is no straightforward way of testing the output of a model

Conclusion

- The development of risk models helps to improve risk awareness and anchors risk management and governance deeper in industry practices
- Risk models provide valuable assessments, especially in relative terms, as well as guidance in business decisions
- It is thus essential to ensure that the results of the model delivers a good description of reality
- Model validation is the way to gain confidence in the model
- It is however difficult because there is no straightforward way of testing the output of a model

Conclusion

- The development of risk models helps to improve risk awareness and anchors risk management and governance deeper in industry practices
- Risk models provide valuable assessments, especially in relative terms, as well as guidance in business decisions
- It is thus essential to ensure that the results of the model delivers a good description of reality
- Model validation is the way to gain confidence in the model
- It is however difficult because there is no straightforward way of testing the output of a model

Conclusion

- The development of risk models helps to improve risk awareness and anchors risk management and governance deeper in industry practices
- Risk models provide valuable assessments, especially in relative terms, as well as guidance in business decisions
- It is thus essential to ensure that the results of the model delivers a good description of reality
- Model validation is the way to gain confidence in the model
- It is however difficult because there is no straightforward way of testing the output of a model

Conclusion

- The development of risk models helps to improve risk awareness and anchors risk management and governance deeper in industry practices
- Risk models provide valuable assessments, especially in relative terms, as well as guidance in business decisions
- It is thus essential to ensure that the results of the model delivers a good description of reality
- Model validation is the way to gain confidence in the model
- It is however difficult because there is no straightforward way of testing the output of a model

Validating a risk model requires the use of various strategies :

- Ensure a good calibration of the model through various statistical techniques
- Use data to test statistically certain parts of the model (like the computation of the risk measure, or some particular model like ESG or Reserving Risk)
- Test the P&L attribution to LoB's against real outcome
- Test the sensitivity of the model to crucial parameters
- Compare the model output to stress scenarios
- Compare the real outcome to its predicted probability by the model
- Examine the simulation output to check the quality of the bankruptcy scenarios (reverse backtest)

Validating a risk model requires the use of various strategies :

- Ensure a good calibration of the model through various statistical techniques
- Use data to test statistically certain parts of the model (like the computation of the risk measure, or some particular model like ESG or Reserving Risk)
- Test the P&L attribution to LoB's against real outcome
- Test the sensitivity of the model to crucial parameters
- Compare the model output to stress scenarios
- Compare the real outcome to its predicted probability by the model
- Examine the simulation output to check the quality of the bankruptcy scenarios (reverse backtest)

Validating a risk model requires the use of various strategies :

- Ensure a good calibration of the model through various statistical techniques
- Use data to test statistically certain parts of the model (like the computation of the risk measure, or some particular model like ESG or Reserving Risk)
- Test the P&L attribution to LoB's against real outcome
- Test the sensitivity of the model to crucial parameters
- Compare the model output to stress scenarios
- Compare the real outcome to its predicted probability by the model
- Examine the simulation output to check the quality of the bankruptcy scenarios (reverse backtest)

Validating a risk model requires the use of various strategies :

- Ensure a good calibration of the model through various statistical techniques
- Use data to test statistically certain parts of the model (like the computation of the risk measure, or some particular model like ESG or Reserving Risk)
- Test the P&L attribution to LoB's against real outcome
- Test the sensitivity of the model to crucial parameters
- Compare the model output to stress scenarios
- Compare the real outcome to its predicted probability by the model
- Examine the simulation output to check the quality of the bankruptcy scenarios (reverse backtest)

Validating a risk model requires the use of various strategies :

- Ensure a good calibration of the model through various statistical techniques
- Use data to test statistically certain parts of the model (like the computation of the risk measure, or some particular model like ESG or Reserving Risk)
- Test the P&L attribution to LoB's against real outcome
- Test the sensitivity of the model to crucial parameters
- Compare the model output to stress scenarios
- Compare the real outcome to its predicted probability by the model
- Examine the simulation output to check the quality of the bankruptcy scenarios (reverse backtest)

Validating a risk model requires the use of various strategies :

- Ensure a good calibration of the model through various statistical techniques
- Use data to test statistically certain parts of the model (like the computation of the risk measure, or some particular model like ESG or Reserving Risk)
- Test the P&L attribution to LoB's against real outcome
- Test the sensitivity of the model to crucial parameters
- Compare the model output to stress scenarios
- Compare the real outcome to its predicted probability by the model
- Examine the simulation output to check the quality of the bankruptcy scenarios (reverse backtest)

Validating a risk model requires the use of various strategies :

- Ensure a good calibration of the model through various statistical techniques
- Use data to test statistically certain parts of the model (like the computation of the risk measure, or some particular model like ESG or Reserving Risk)
- Test the P&L attribution to LoB's against real outcome
- Test the sensitivity of the model to crucial parameters
- Compare the model output to stress scenarios
- Compare the real outcome to its predicted probability by the model
- Examine the simulation output to check the quality of the bankruptcy scenarios (reverse backtest)