UBS

Public

Model Validation

Workshop

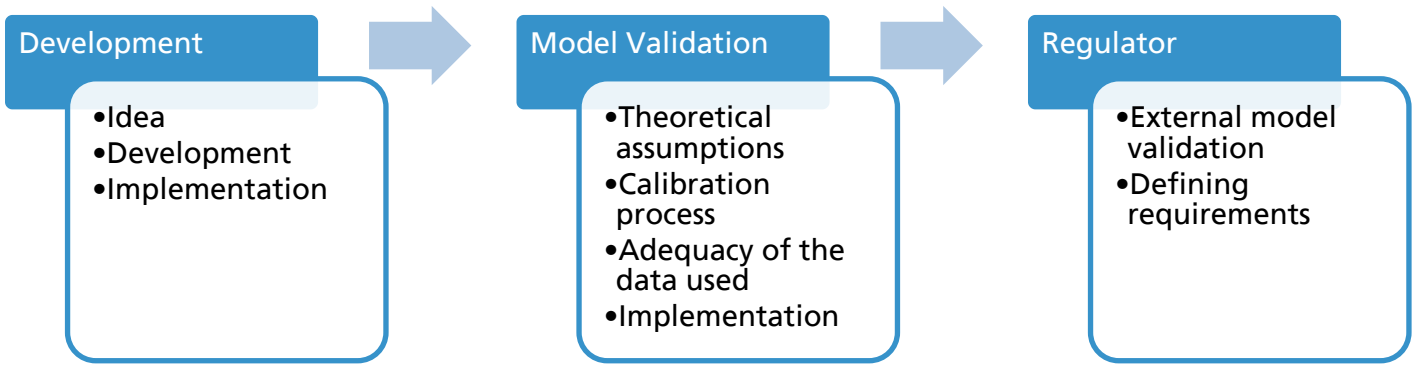
dr Piotr Morawski
Adam Wróbel
Model Risk Management & Control

May 17, 2016

Model Validation

What is this all about?

- Enforced on financial institutions by regulators
- With aim to ensure that whenever bank use model it is done in appropriate manner
- Especially relevant for:
  - capital requirement related models
  - used in trading models (front desk)



Credit Risk

What is the credit risk?

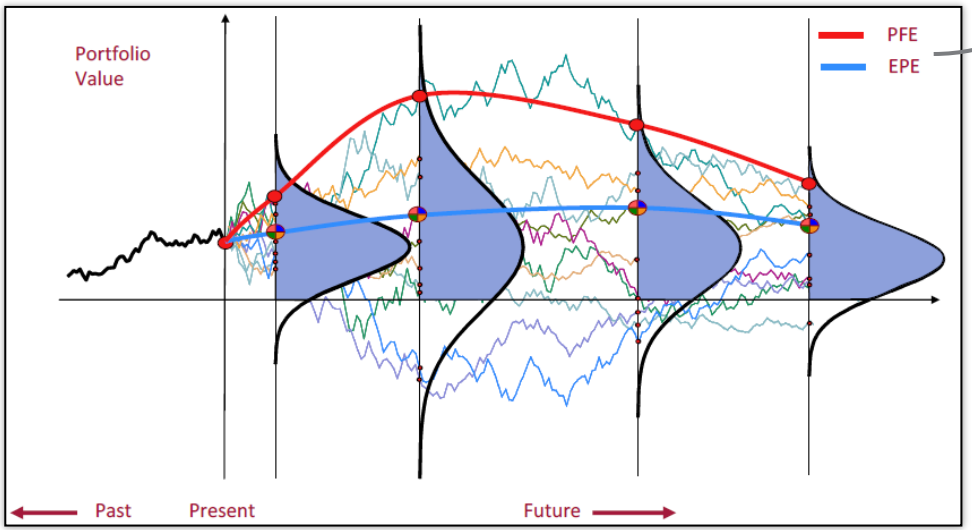
I. Risk of the counterparty defaulting:
   - Client unable to pay the mortgage
   - Friend unwilling to pay us back
   - Lehman Brothers couldn't keep their end of the bargains

II. Depends on various factors
   - Quality of the counterparty – Friend with stable income is less risky one
   - Economy - crisis could decrease company ability to sell its product
   - Management – company even in very bad financial situation could improve instead of going into the default

III. How to quantify credit risk?
   - PD – Probability of Default
   - LGD – Loss Given Default
   - EAD – Exposure at Default
   - Expected Loss = PD * LGD * EAD

Credit Risk

Credit Risk Pricing



courtesy of Giovanni Cesari, UBS

- $\mathrm{CVA} = \mathrm{LGD} \times \int_0^T EAD \times PD \; du$

Exposure At Default

Loss Given Default

Probability of Default

- Pricing the credit risk requires three ingredients:
  - LGD, PD – counterparty dependent
  - EAD – derivative dependent

- Exposure profilling:
  - Calculating CVA requires knowledge of Trade Value distribution at each point in time
  - MC simulation can provide very general solution to problem of pricing a derivatives regardless of it's complexity
- LGD, PD modelling:
  - For CVA PD/LGD comes from market implied spread
  - For other purposes statistical models based on internal or external historical data

R fundamentals

During the workshop we will extensively use R

We will mainly use base functions and *dplyr* approach for data processing, but obviously you are free to use whatever fit you best during the exercises.

Let's open script "*0_R_introduction.R*" and get familiar with some functions that we will use

Problem 1

Normally distributed variable?

- Most popular assumption done during the modelling
- Multiple well know tests could be applied, but should one blindly take test result without looking into the data?
- Open script "*1_Tests.R*" and let's take a closer look

Problem 2

Linear regression used for modelling default probability

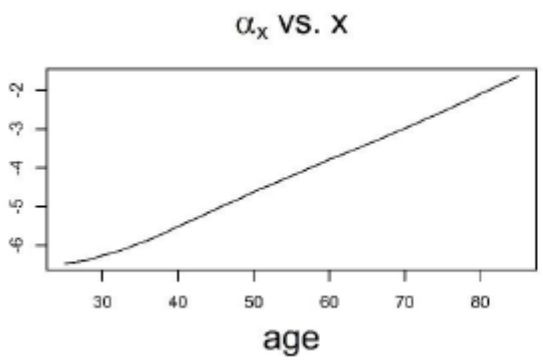
- Some bank offers short term loans with relatively high risk of clients not fulfilling their obligations
- Data set* with three variables:
  - *income* – monthly income of a client
  - *loan_size* – amount of money that client own to the bank
  - default – binary variable saying if client did default or not
- Model defined as: $default \sim intercept + \beta_1 income + \beta_2 loan\ size + error\ term$
- Let's open script "*2_Linear_regression.R*" and check if this approach is appropriate

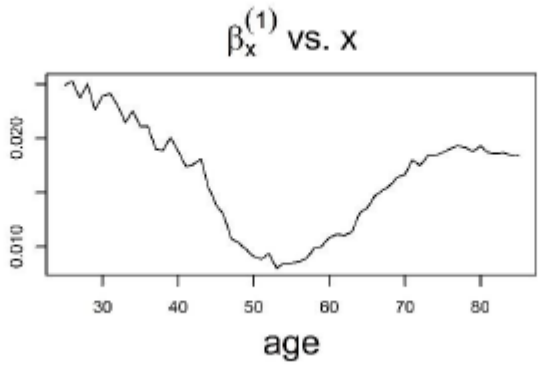
* all the datasets used during this workshop are simulated or publicly available

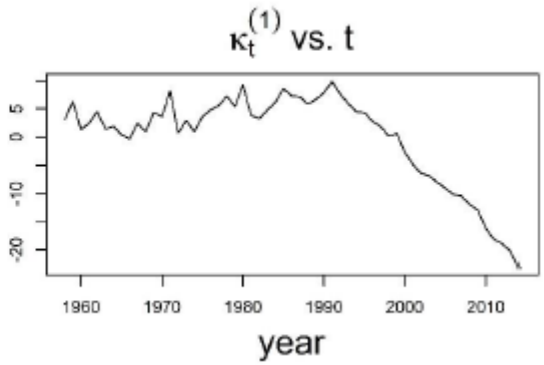
Problem 3

Population death probability modelled by Lee-Carter model *("3_stochastic_mortality.R")*

- With aim to model death probability of male population of Poland (based on mortality tables)
- Generalized Linear Model (GLM)
- Number of deaths as modelled variable (assumed binomial distribution)
- Logit link function (same as in logistic regression)
- Lee-Carter specifically is defined as: $\eta_{xt} = \alpha_x + B_x^{(1)} \kappa_t^{(1)}$



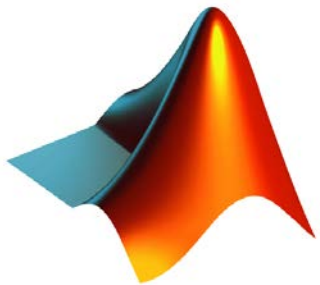




Tools

What are the main tools used in the industry? What should I learn?

- Prototyping the models is most often done in R, SAS, Matlab
- Production version of models are often written in C++ and C#
- Some models (often those that are used for staff like capital requirements calculation, stress testing) have production implementation in R, SAS, Matlab
- Other environments/languages that are also popular are Python and VBA
- Where is the future:
  - R
  - Python




Problem 4

Testing model's prediction power on historical data

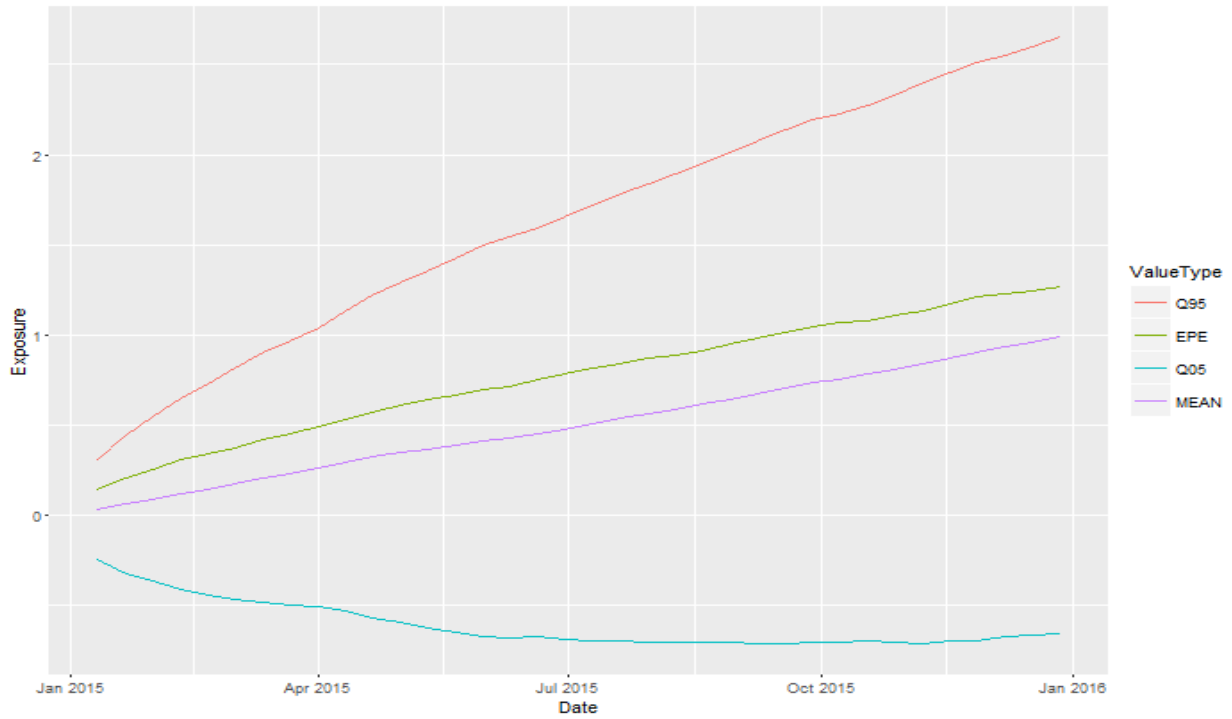
- We need to model the trade value throughout the whole trade period to know how much we are exposed in case of the counterparty gone bankrupt.
- To do that we need to simulate the Risk Factors driving the trade value.
- In simple case of stock option the only Risk factor relevant is the price of the underlying stock.
- Model: linear model calibrated over the last month of data, used to predict the next day stock price.
  - Underlying assumption: the stock prices is a Black-Sholes process.
- **Principle of backtesting**: using available market data to compare model prediction against the historical realisation.
- Use the script: "*4_Backtesting.R*"

UBS

Problem 5

CVA calculation

- Profile the exposure of a simple stock call option
  - Simulate underlying stock movement using MC technique
  - Measure the Expected Positive Exposure at each point in time
- Combine the Exposure profile with Default Probability and Loss Given Default of the counterparty.
  - Open script: 5_CVA.R



People

Who works in risk modelling/quantitative finance?

- Most common background:
  - Mathematics
  - Econometrics / Quantitative finance
  - Physics
- PhD's

Problem 6

Loss Given Default (LGD) model – building benchmark model

- It is relevant not only if counterparty would default, but also the percentage of our loan that we could lost in case od default
- Dataset: US clients with collateralized loans that did default at some point in time (also macroeconomic variables at the year of default are also included in the dataset)
- Based on historical default data probit model was fitted:
  $loss\ given\ default \sim intercept + \beta_1 loan\ to\ value + \beta_2 GPD\ growth + \beta_3 unemplyment\ rate + \beta_4 inflation +\ error\ term$
- Let's open script "*6_LGD.R*"

Problem 7

Stress model – are all variables stressed in appropriate manner?

- We would like to know how much we could lost if another crisis would hit us (alternatively any other scenario)
- One approach to model the loss is to use build PD and LGD models and then stress their inputs:
  - increased unemployment rate
  - decreased growth of gross domestic product
  - drop of the house prices
- Let's open script "*7_Stress_model.R*"

Sum up

What are the most important takeaways?

- Appling model without checking it could result in an underestimation of risk or required capital buffer.
- Which could create incentive to get into more risky transitions. On single bank perspective it could lead to huge losses and potential bankruptcy. While on a whole financial system level it could result in crisis.
- There is a huge universe of models and if you enjoy getting to understand every one of them then model validation is an areas that is worth considering.

Contact information

dr Piotr Morawski
UBS Krakow, Risk Control Specialist
Tel. +48 12 399 81 79
piotr-a.morawski@ubs.com

Adam Wróbel
UBS Krakow, Risk Modelling & Analytics Specialist
Tel. + 48 12 399 69 65
adam.wrobel@ubs.com

ubs.com/polandcareers