

# Risk Modeling

Within banking & insurance

Adam Wróbel

June 16, 2018

# Agenda

1. Risk modeling
  - a) In banking
  - b) In insurance
2. Model development & modal validation
3. R in risk modelling
4. Practical examples:
  - a) Risk aggregation
  - b) Stochastic mortality

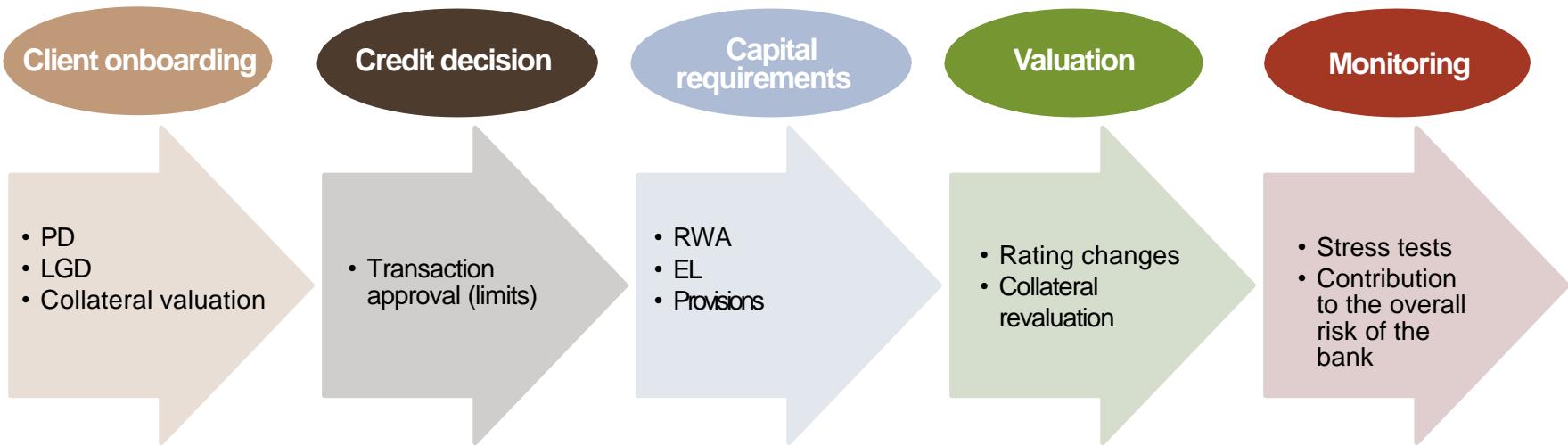
# Risk modeling in banking industry

## 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

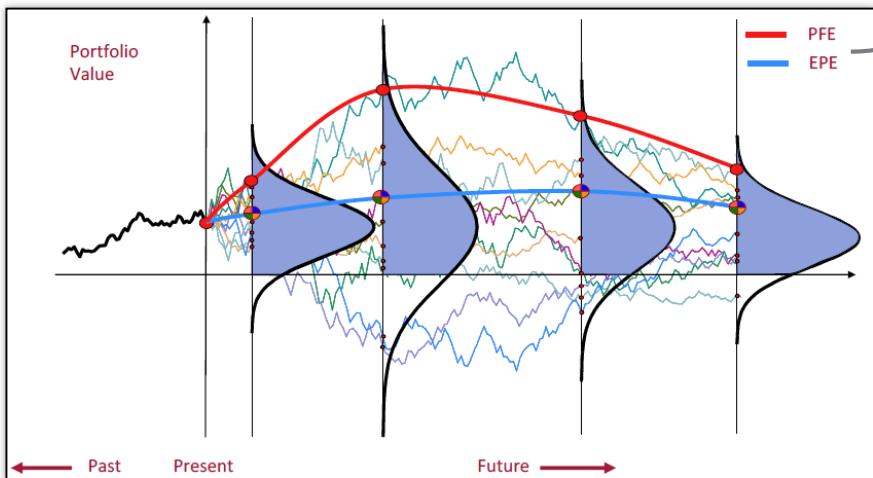
# Risk modeling in banking industry

## Credit decision example



# Risk modeling in banking industry

## Credit Risk Pricing



- $CVA = LGD \times \int_0^T EAD \times PD \, du$

Loss Given Default      Exposure At Default      Probability of Default

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

courtesy of Giovanni Cesari, UBS

- 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

# Risk modeling in insurance industry

## Examples

- Life insurance
  - Modelling death probability
  - Probability of withdrawing the policy
  - Pricing of the life insurance policies & riders
  - Long term cash flow modeling (up 20-30 years Ahead)
  - Replicating portfolio
- Non-life insurance:
  - Pricing for all segments (automotive, housing)

# Both in banking & insurance

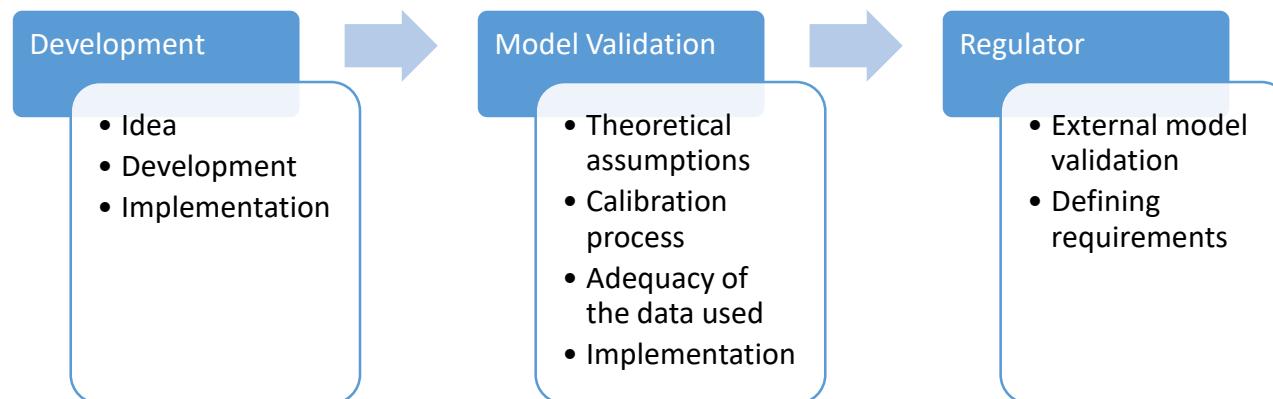
## Examples

- Market Risk:
  - Valuation models
  - Value at Risk
  - Economic scenario generation
- Consequential Risks:
  - Operational Risk
  - Business Risk (risk coming from income volatility)
  - Funding Risk

# Model Validation

What is this all about?

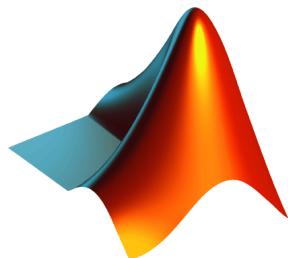
- Enforced on financial institutions by regulators
- With aim to ensure that whenever bank use model it is done in appropriate manner
- Especially relevant for:
  - capital requirement related models
  - used in trading models (front desk)



# 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



# Agregacja ryzyka

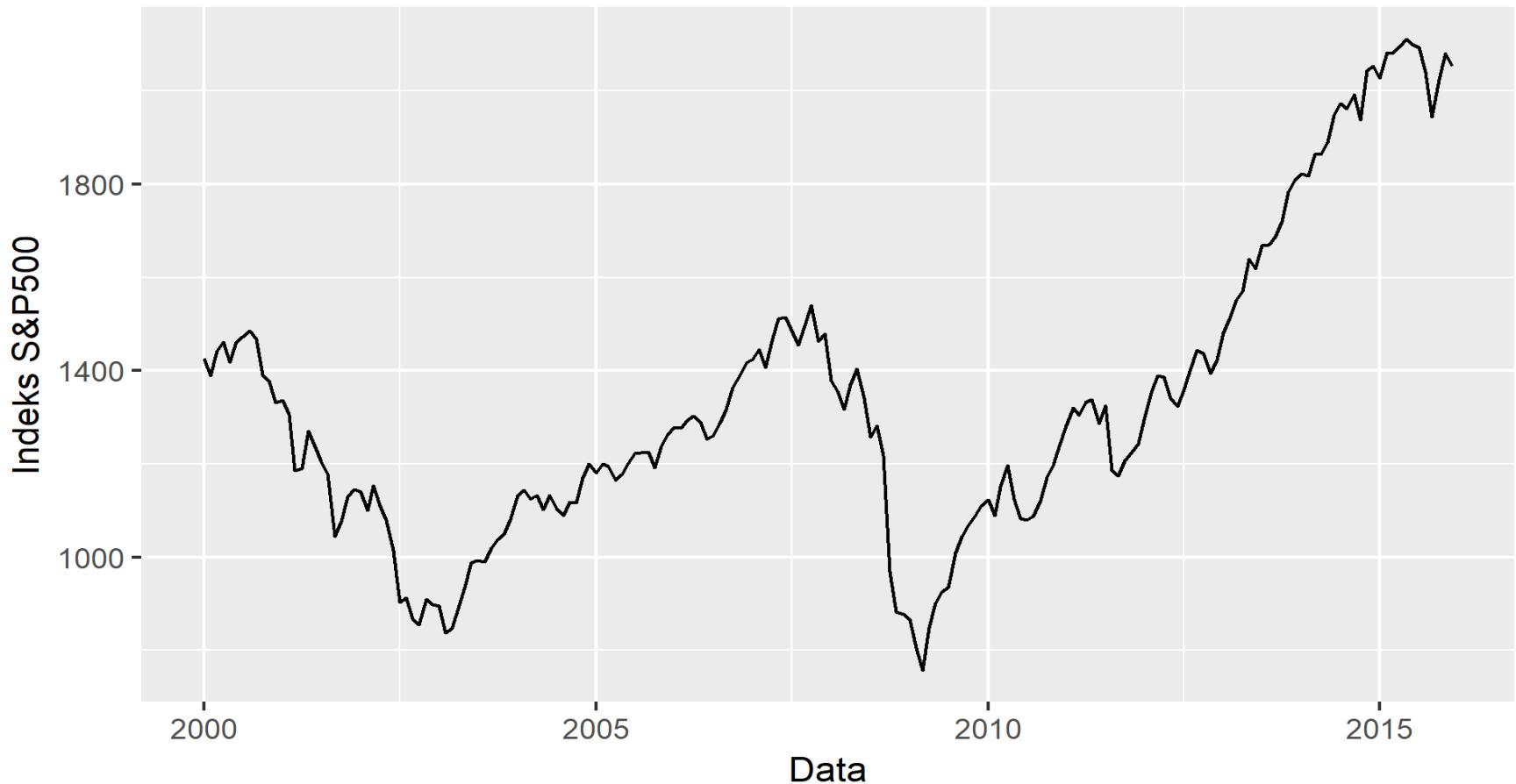
Zagrożenia w perspektywie działalności banków

Banki są narażone na wiele czynników ryzyka:

- krach na giełdzie
- spadek cen nieruchomości
- pogorszenie sytuacji gospodarczej
- zmiany kursu walutowego

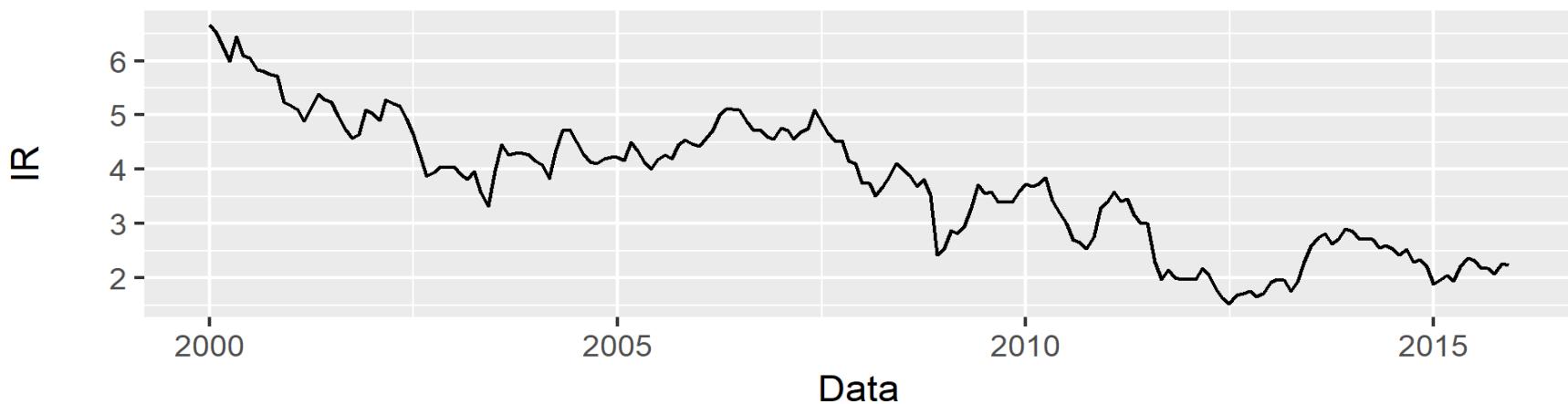
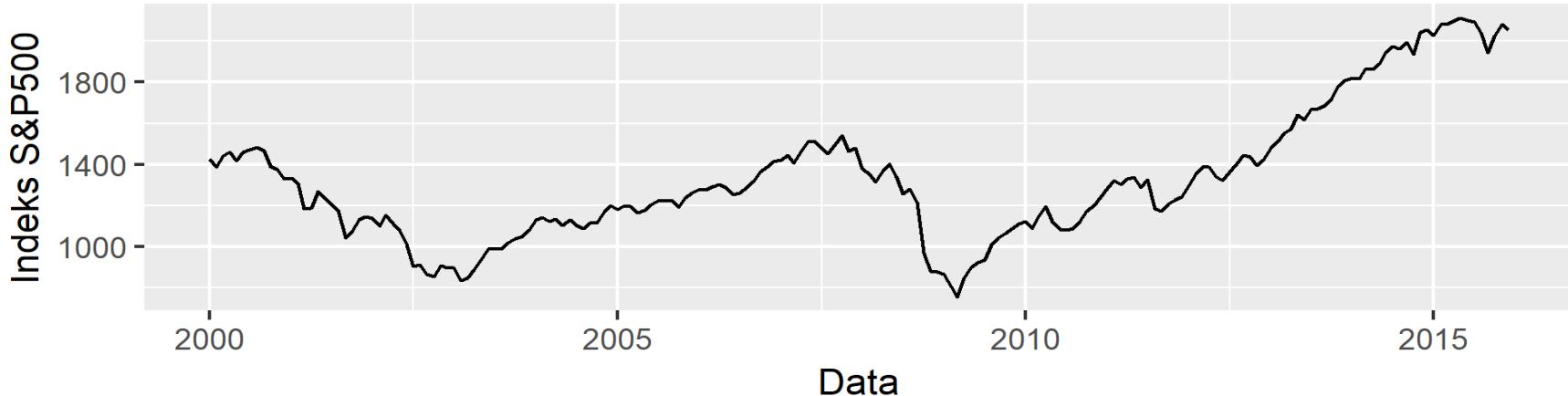
# Sytuacja na giełdzie - rynki finansowe

Notowania indeksu S&P500 – index 500 największych spółek na amerykańskiej giełdzie



# Zależność między czynnikami ryzyka

Indeks giełdowy S&P500 oraz stopy procentowe (IR – interest rate)



# Modelowanie rozkładu strat/zysków

Jak może zmienić się wartość naszych pozycji (akcji, obligacji) w ciągu kwartału?

- Cała ekspozycja to akcje spółek notowanych na amerykańskiej giełdzie oraz obligacje rządu USA
- W uproszczeniu możemy przyjąć, że:
  - zmiana wartości danej akcji będzie zależała od zmiany wartości indeksu S&P500
  - zmiana wartości danej obligacji będzie zależała od zmiany stopy procentowej
- Mając modele wyceny (patrz punkt wyżej) możemy wyznaczyć zmianę wartości naszych pozycji (akcji i obligacji) przy danej realizacji czynników ryzyka
- Mając historyczne dane o zmienności czynników ryzyka możemy wysymulować realizacje tych dwóch czynników z uwzględnieniem zależności pomiędzy nimi
- Mając odpowiednio dużo tychże realizacji możemy wyznaczyć rozkład strat/zysków wyceniając wartość pozycji w każdym scenariuszu

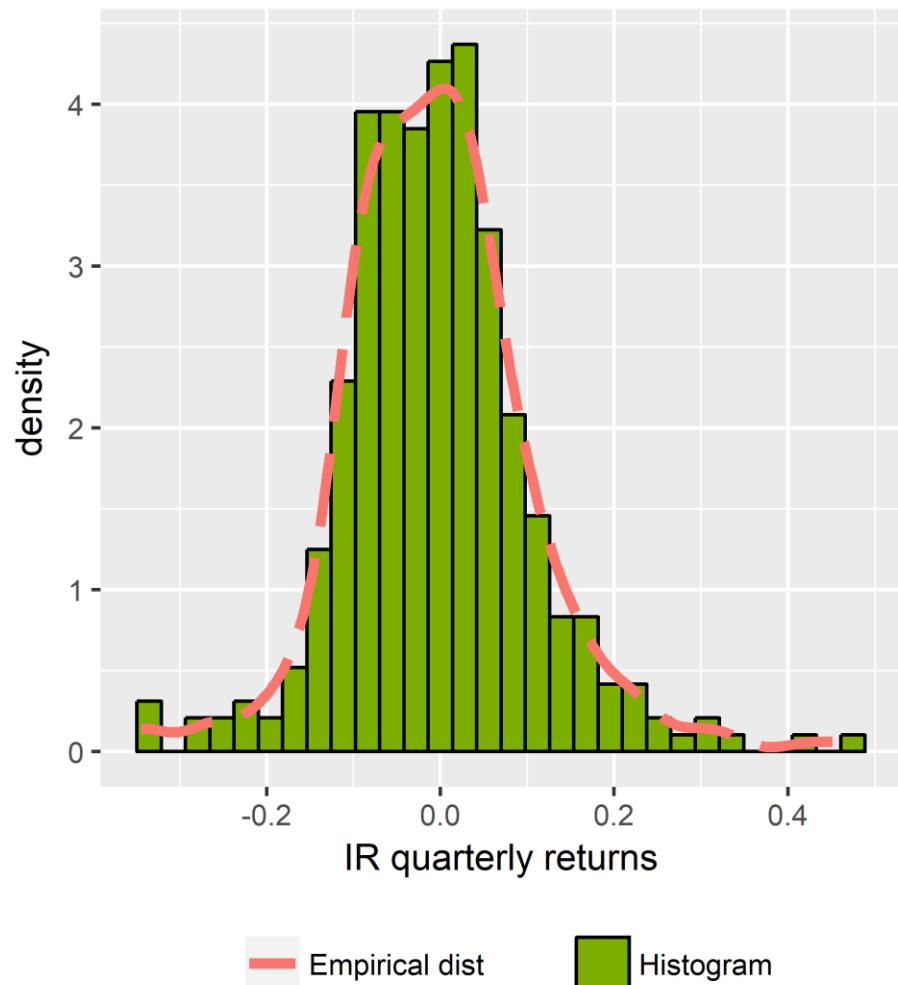
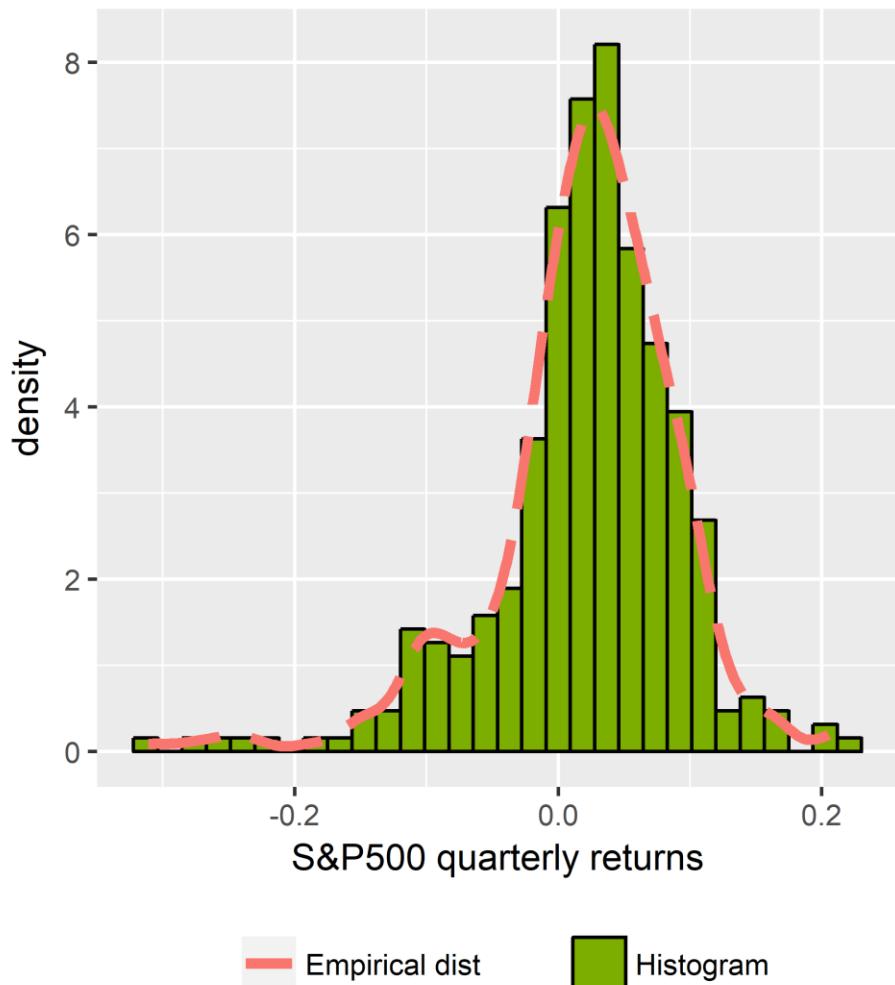
# Modelowanie rozkładu strat/zysków

Jak może zmienić się wartość naszych pozycji (akcji, obligacji) w ciągu kwartału?

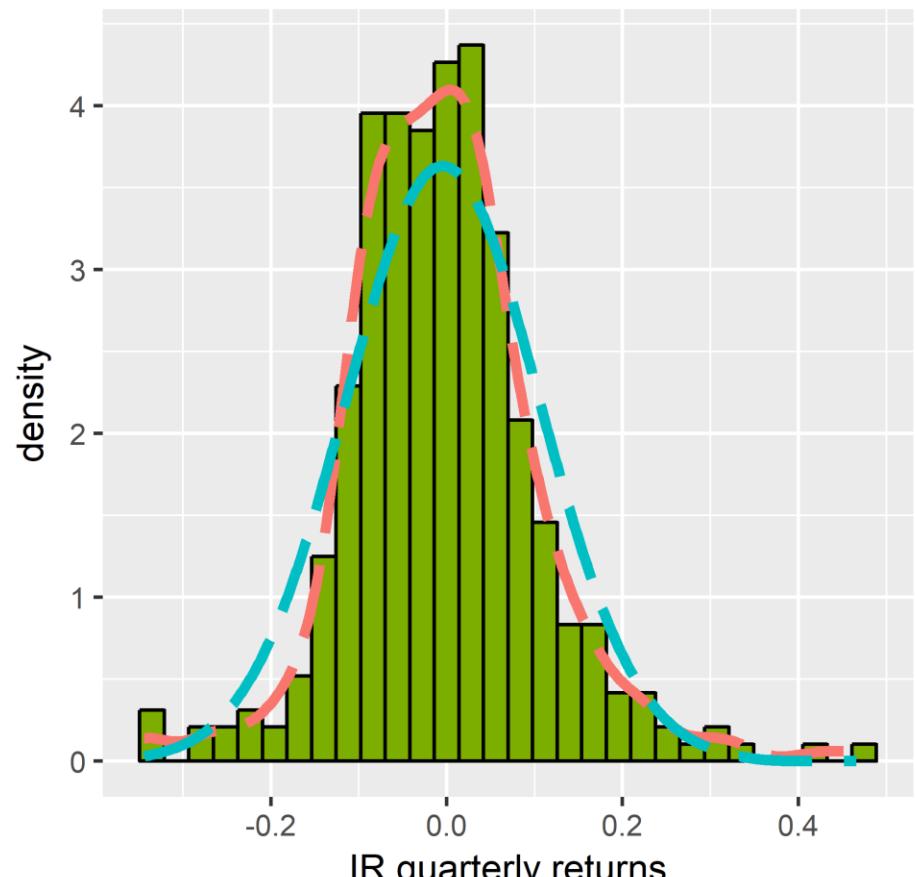
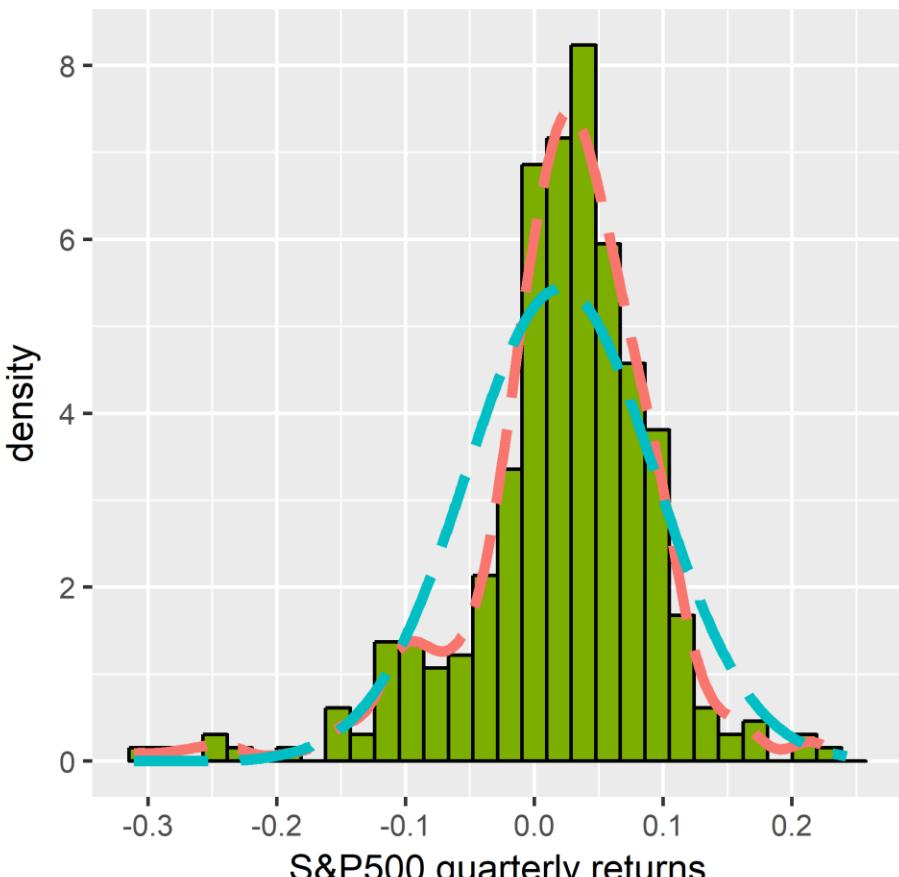
Scenariusz	$\Delta$ S&P500	$\Delta$ stopy procentowej	zmiana wartości akcji A	zmiana wartości obligacji B	Strata/Zysk
1	5%	0%	20 000\$	0\$	20 000 \$
2	-5%	9%	-20 000\$	-7000\$	-27 000 \$

- $\Delta$  – kwartalna relatywna zmiana czynnika (simple return):  $\Delta x_t = \frac{x_t - x_{t-1}}{x_{t-1}}$
- Mając N takich scenariuszy otrzymujemy rozkład strat/zysków

# Rozkłady kwartalnych zmian S&P500 i IR



# Dopasowane rozkłady normalne



Empirical dist    Fitted normal dist

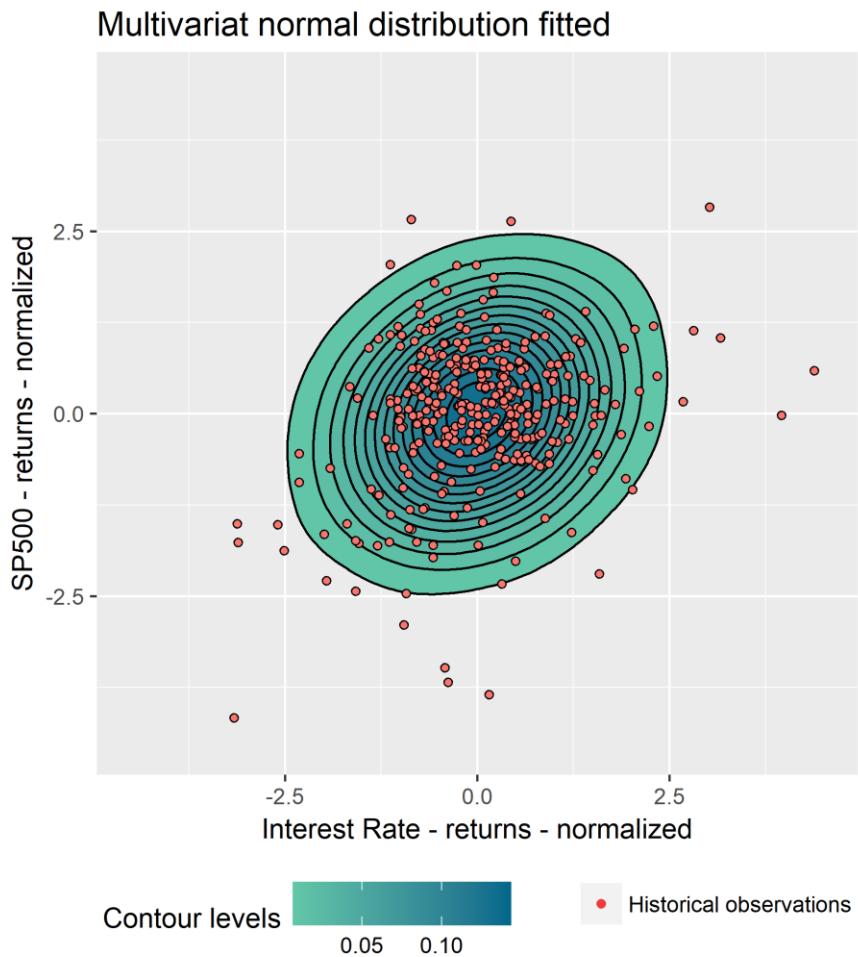
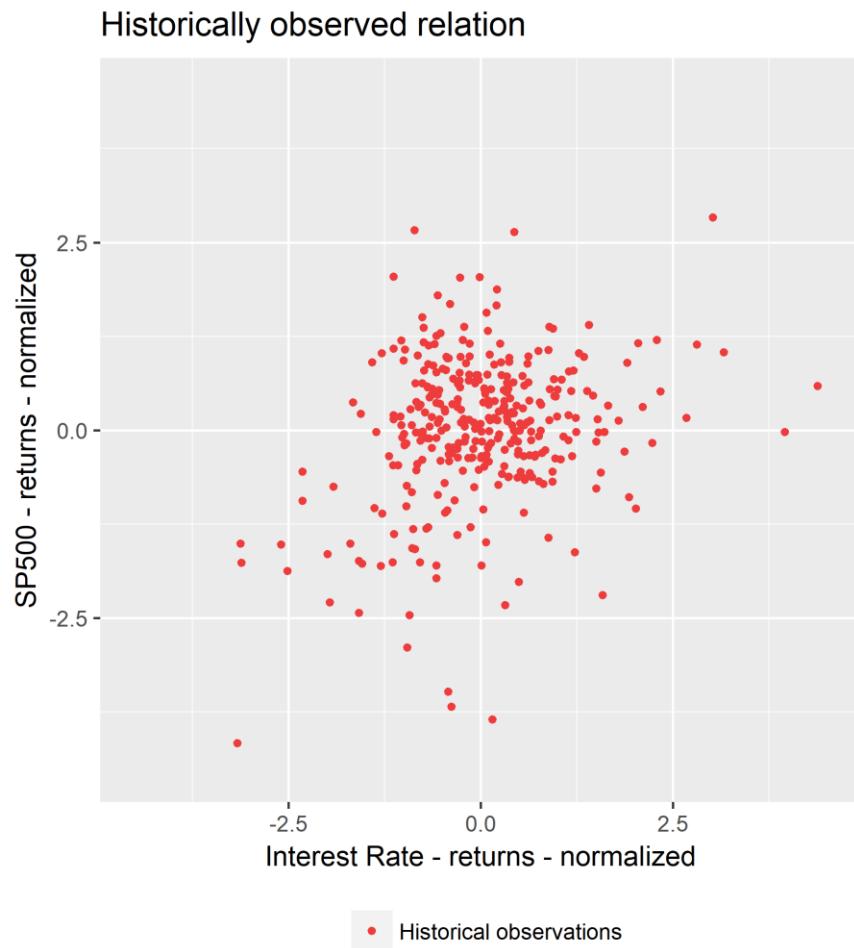
Histogram

Empirical dist    Fitted normal dist

Histogram

# Zależność pomiędzy rozkładami

Próba modelowania zależności przy pomocy wielowymiarowego rozkładu normalnego



# Wprowadzenie do kopuł

- Wielowymiarowy rozkład normalny jest kopułą Gaussa z normalnymi rozkładami brzegowymi
- Kopuły pozwalają rozdzielić modelowanie na dwa niezależne kroki:
  - Zdefiniowanie rozkładów brzegowych
  - Zdefiniowanie zależności między rozkładami brzegowymi
- Kopuła Gaussa wykorzystuje tą samą strukturę zależności jak w wielowymiarowym rozkładzie normalnym. Czyli potrzebujemy jedynie macierzy korelacji.
- Z matematycznego punktu widzenia zawsze istnieje kopuła łącząca dane rozkłady brzegowe. Jeśli rozkłady brzegowe są ciągłe to jest też to unikalna kopuła (Sklar's theorem).

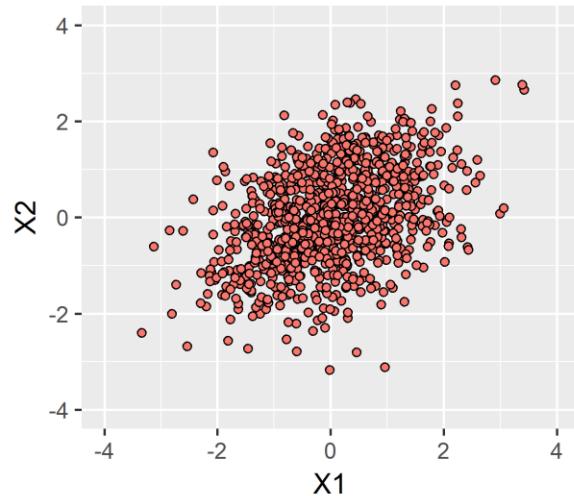
# Kopuły

## Budowanie kopuły

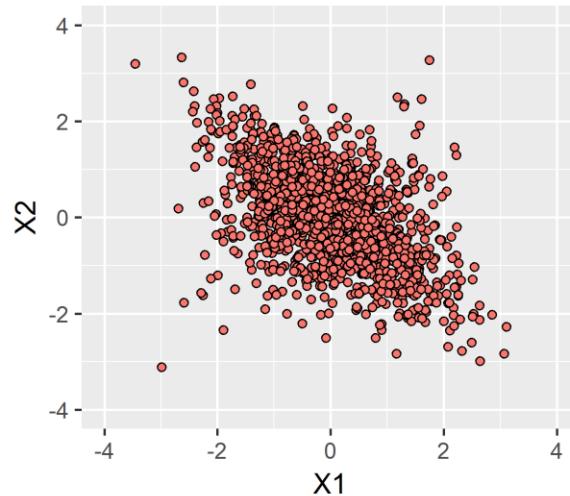
- Kopuła definiuje zależność pomiędzy rozkładami jednostajnymi:
    - mając dopasowany rozkład teoretyczny do danych możemy zawsze przetransformować go do jednostajnego
  - Kroki budowy modelu:
    - Dopasowanie rozkładów teoretycznych do rozkładów brzegowych (pakiety: *fitdistrplus*, *ghyp*)
    - Dopasowanie kopuły (pakiety: *CDVine*, *VineCopula*)
  - Przykład. Kroki symulacji z kopuły Gaussa z rozkładami brzegowymi t studenta:
    - Symulacja rozkładów brzegowych z dopasowanych rozkładów teoretycznych
    - Transformacja do rozkładów jednostajnych
    - Uzależnienie rozkładów brzegowych wykorzystując macierz korelacji
    - Transformacja do rozkładów jednostajnych
    - Transformacja do dopasowanych na samym początku rozkładów brzegowych
- ```
> rt %>% pnorm %>% qnorm %*% chol(cholesky_decomposed_cor_matrix) %>% pnorm %>% qt
```

# Kopuły - przykłady

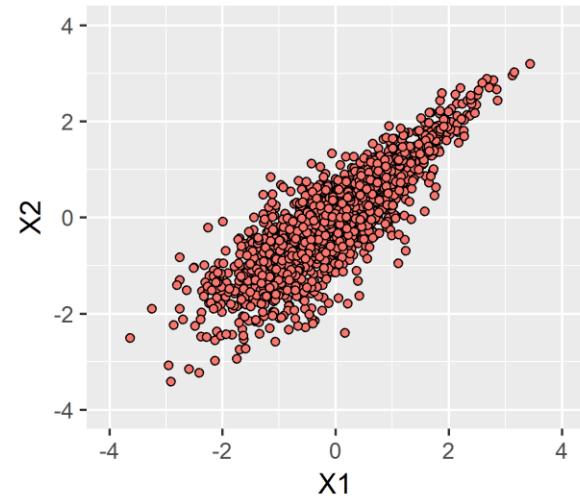
Gaussian Copula - Simulated



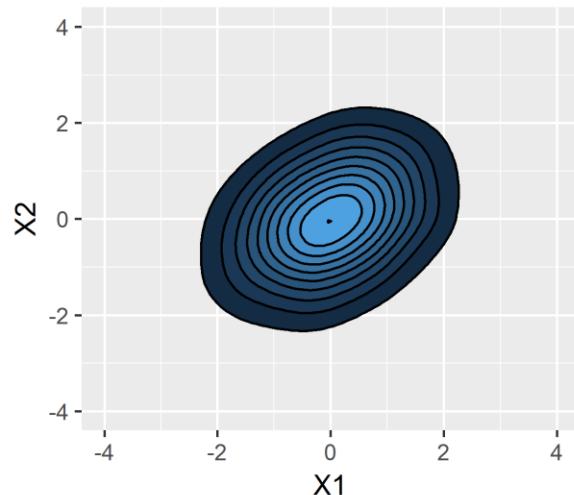
T copula - Simulated



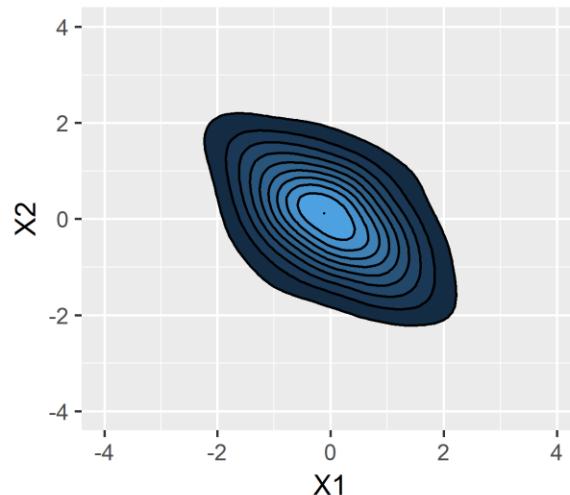
Gumbel copula - Simulated



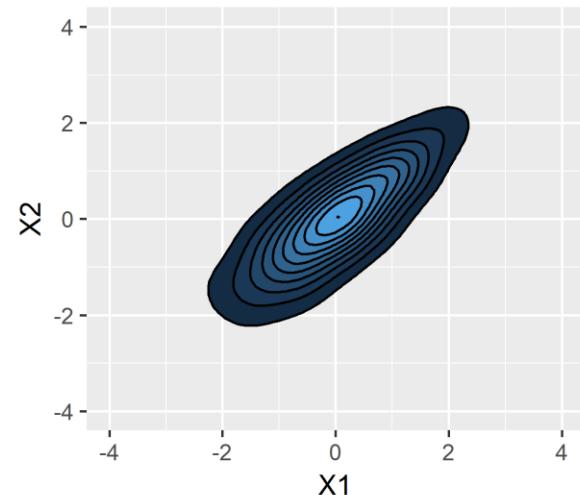
Gaussian Copula - Contour



T copula - Contour



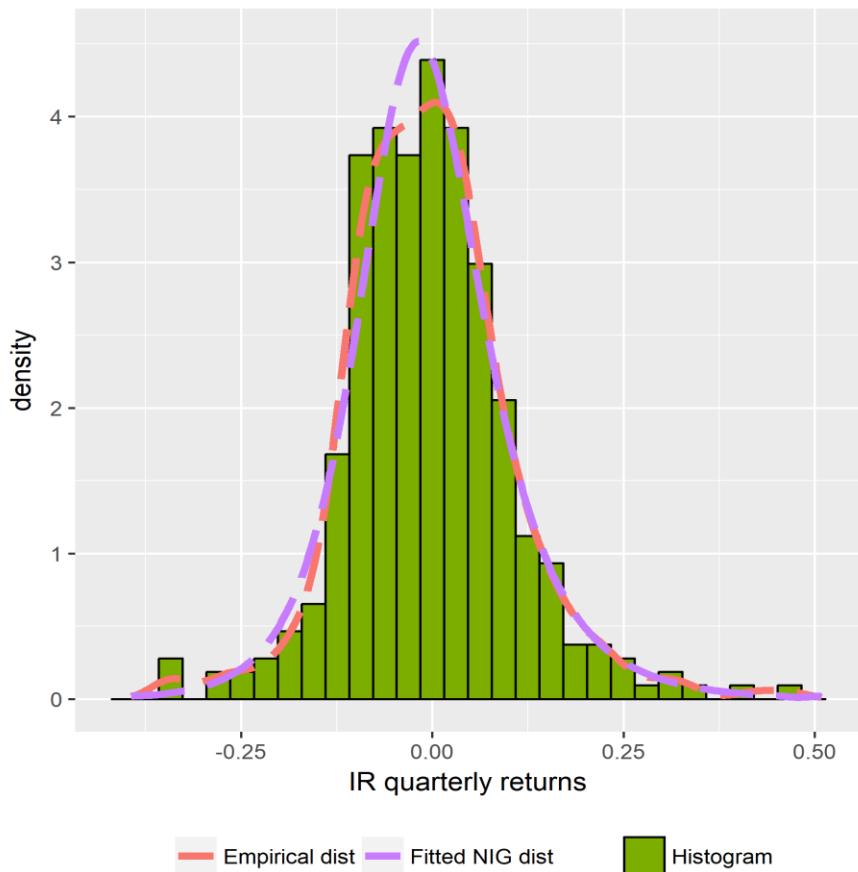
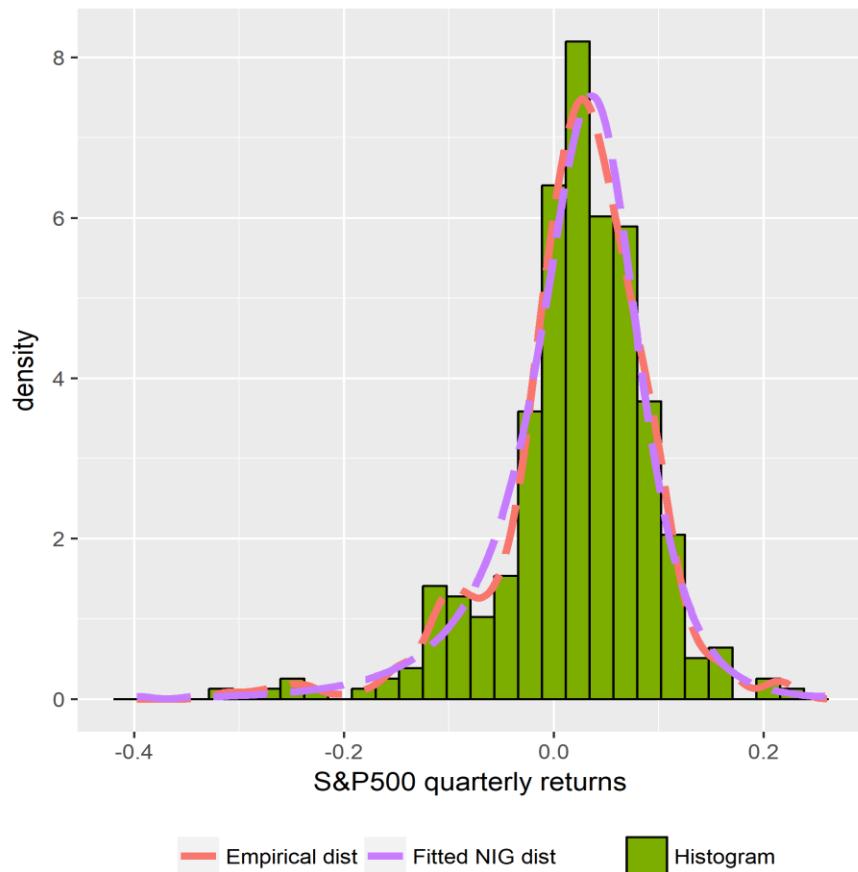
Gumbel copula - Contour



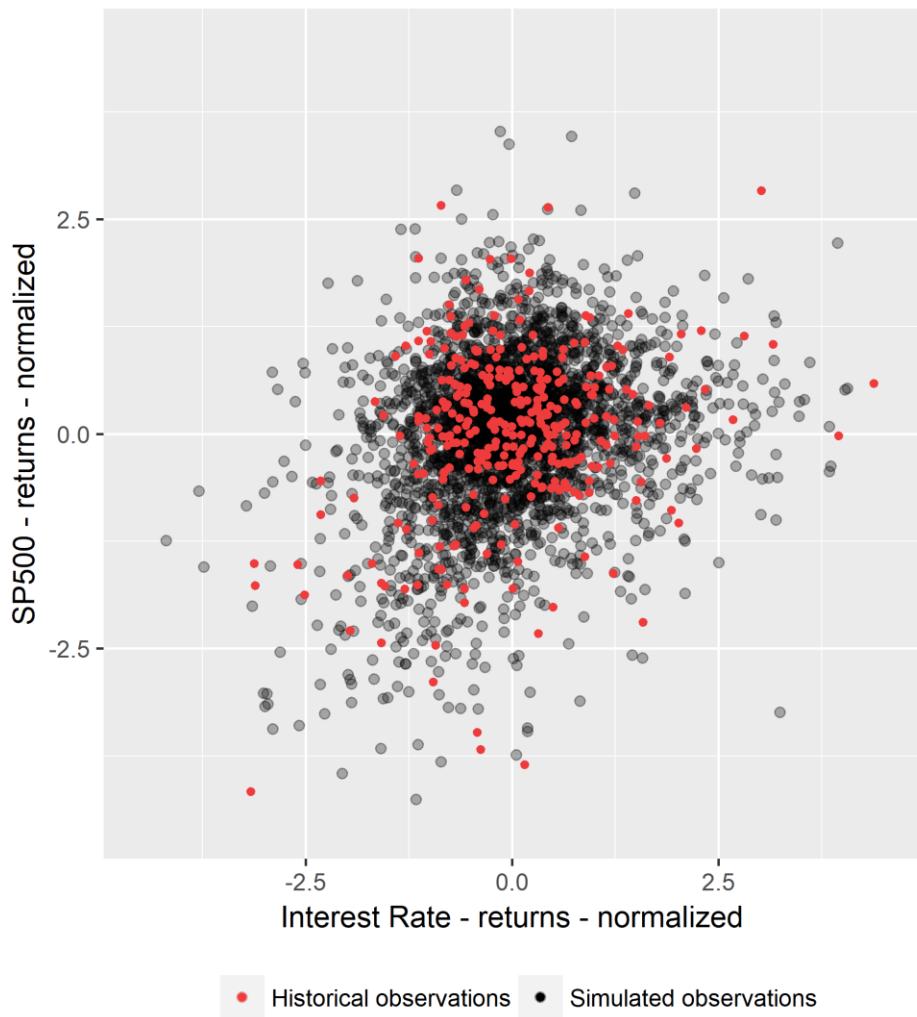
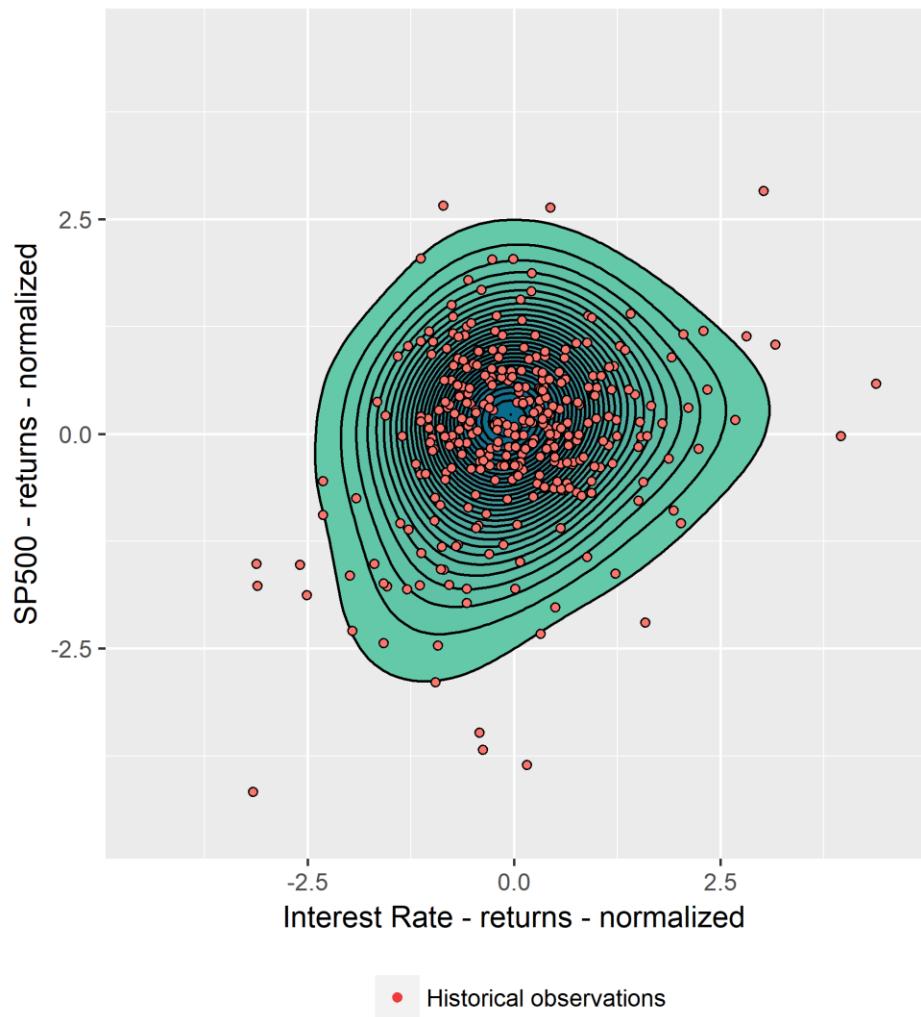
Przyjęte rozkłady brzegowe: standardowy normalny

# Wyniki: Dopasowanie rozkładów

Normal Inverse Gaussian (NIG); pakiet ghyp



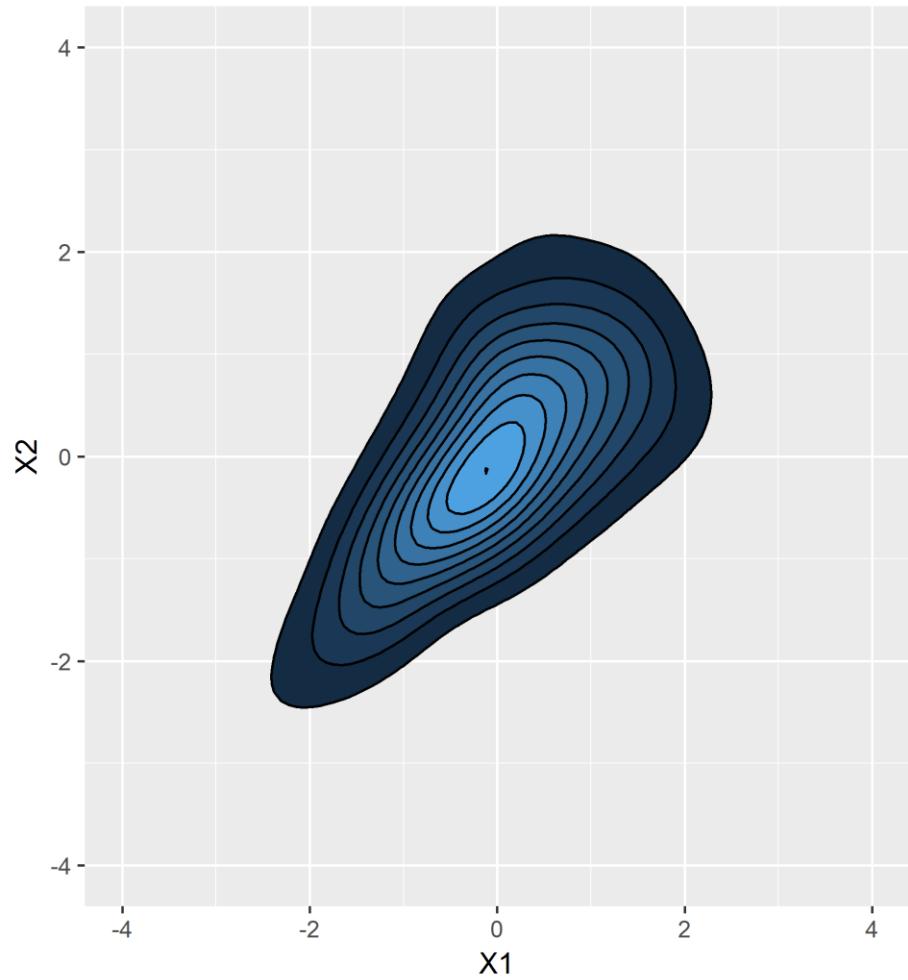
# Wyniki: Porównanie z historycznymi danymi



# Dopasowanie kopuły

Joe copula

Contour



Simulated observations



\*survival version

# Dodatkowe komentarze

- Alternatywa w postaci **historical simulation**:
  - Symulacja z empirycznych rozkładów
  - Historyczna zależność pomiędzy czynnikami ryzyka
- Bardzo szeroko wykorzystywane w modelach wewnętrznych (Internal Models) w ubezpieczeniach razem z **replicating portfolio**

# Polecana literatura/materiały

- U. Schepsmeier, E. C. Brechmann, Modeling dependence with C- and D-vine copulas: The R package CDVine, 2013
- U. Schepsmeier, E. C. Brechmann, CDVine, 2013  
<https://cran.r-project.org/web/packages/CDVine/>
- H. Joe, Dependence Modeling with Copulas, 2014  
(not yet published on CRAN package CopulaModel - [copula.stat.ubc.ca](http://copula.stat.ubc.ca))
- C. Genest, A.-C. Favre, Everything You Always Wanted to Know about Copula but Were Afraid to Ask, 2007
- H. Joe, H. Li, A. K. Nikoloulopoulos, Tail dependence functions and vine copulas, 2010
- M. Hofer, M. Machler, Nested Archimedean Copulas Meet R – Vignette,  
<https://cran.r-project.org/web/packages/copula/vignettes/nacopula-pkg.pdf>

# Polecana literatura/materiały

- U. Schepsmeier, E. C. Brechmann, Modeling dependence with C- and D-vine copulas: The R package CDVine, 2013
- U. Schepsmeier, E. C. Brechmann, CDVine, 2013  
<https://cran.r-project.org/web/packages/CDVine/>
- H. Joe, Dependence Modeling with Copulas, 2014  
(not yet published on CRAN package CopulaModel - [copula.stat.ubc.ca](http://copula.stat.ubc.ca))
- C. Genest, A.-C. Favre, Everything You Always Wanted to Know about Copula but Were Afraid to Ask, 2007
- H. Joe, H. Li, A. K. Nikoloulopoulos, Tail dependence functions and vine copulas, 2010
- M. Hofer, M. Machler, Nested Archimedean Copulas Meet R – Vignette,  
<https://cran.r-project.org/web/packages/copula/vignettes/nacopula-pkg.pdf>

# Stochastic Mortality

Since we have got data on number of deaths and size of a population we are able to calculate death probability. We will focus on conditional death probability within one year (so called  $q_x$ )

Factors that we will take into considerations are as fallow:

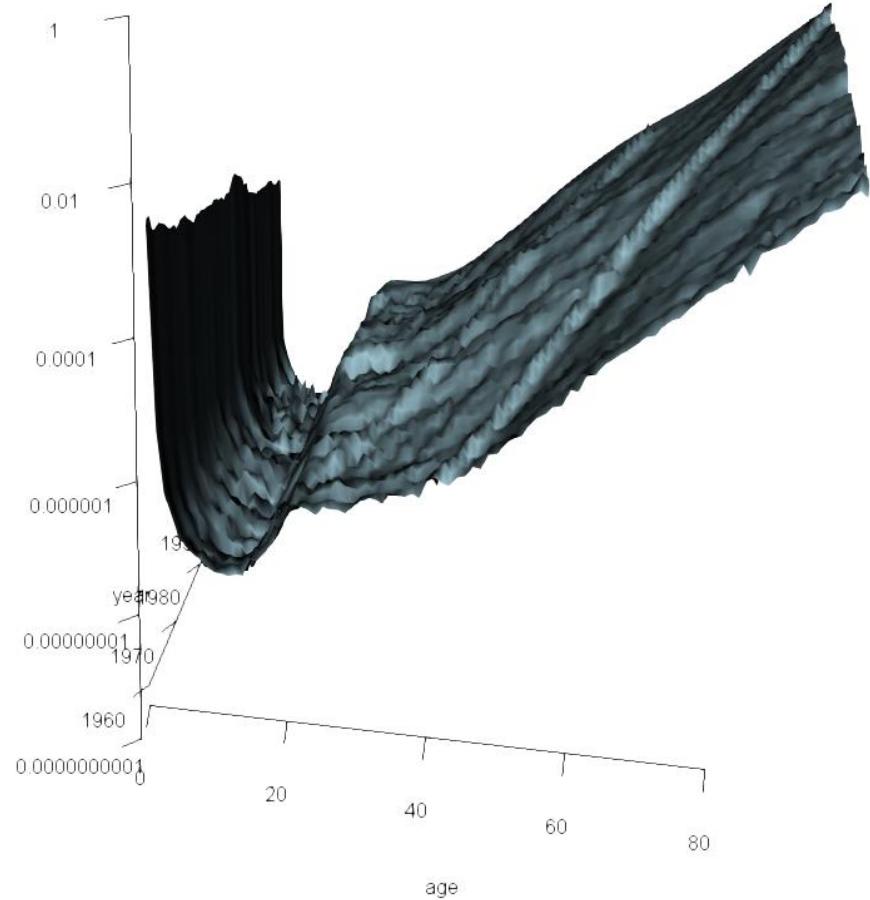
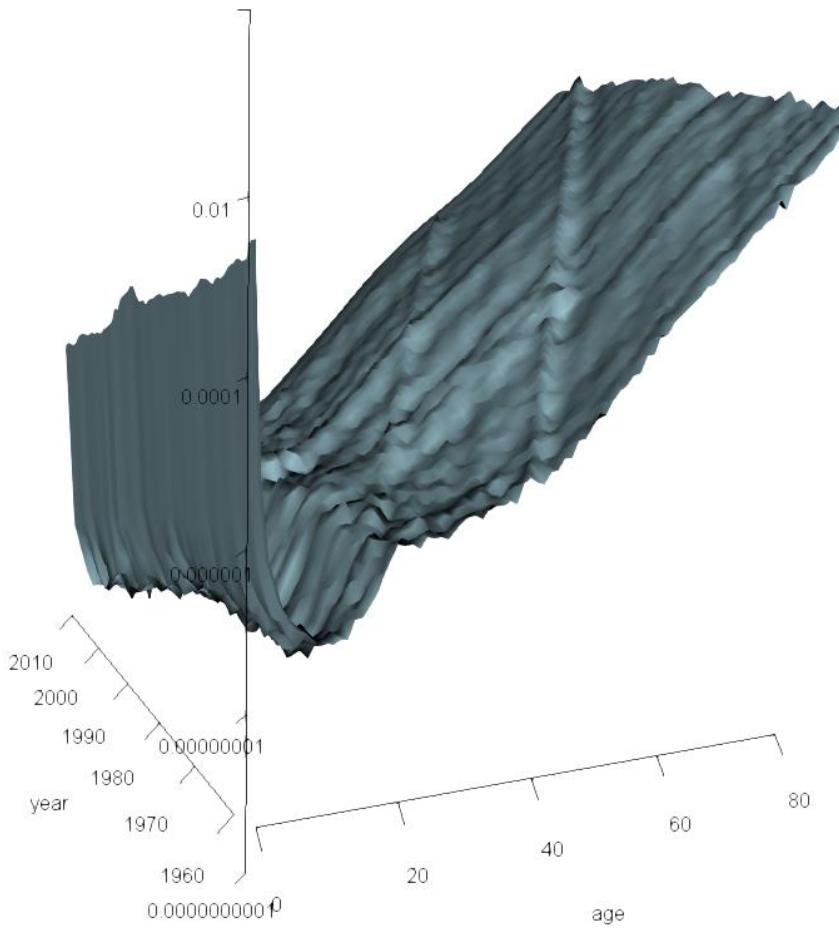
- Gender
- Age
- Calendar Year
- Cohort (year of birth)

# Mortality tables

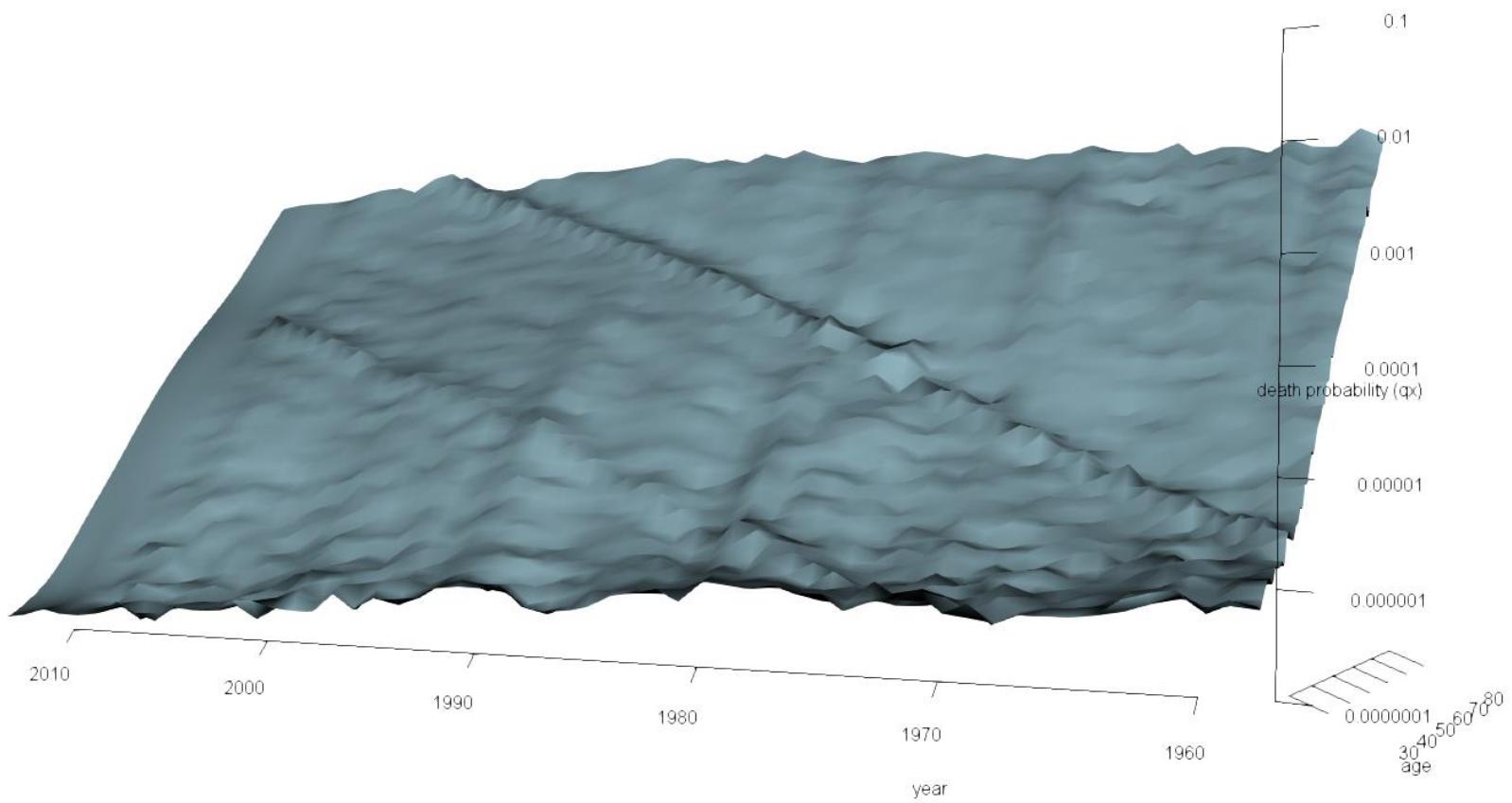
Framework that is used to capture this conditional death probability is called mortality table.

| Age\Year | 2010     | 2011     | 2012     | 2013     | 2014     |
|----------|----------|----------|----------|----------|----------|
| 25       | 0.001047 | 0.001047 | 0.001098 | 0.001016 | 0.000984 |
| 26       | 0.001069 | 0.001058 | 0.001109 | 0.001027 | 0.000995 |
| 27       | 0.001101 | 0.0011   | 0.00111  | 0.001049 | 0.001037 |
| 28       | 0.001153 | 0.001153 | 0.001122 | 0.00108  | 0.001079 |
| 29       | 0.001205 | 0.001225 | 0.001143 | 0.001132 | 0.001131 |
| 30       | 0.001278 | 0.001299 | 0.001196 | 0.001175 | 0.001183 |
| 31       | 0.001372 | 0.001392 | 0.001269 | 0.001248 | 0.001246 |
| 32       | 0.001477 | 0.001507 | 0.001363 | 0.001321 | 0.001319 |
| 33       | 0.001602 | 0.001633 | 0.001488 | 0.001425 | 0.001403 |
| 34       | 0.001739 | 0.00178  | 0.001624 | 0.00154  | 0.001507 |

# Visualization of mortality tables



# Visualization of mortality tables with restrictions



# Data source

**Human Mortality Database (1958-2009):**

<http://www.mortality.org>

**Central Statistical Office of Poland (1990-2014):**

<http://stat.gov.pl/en/topics/population/life-expectancy/life-expectancy-in-poland,1,1.html>

# Stochastic mortality models

Stochastic mortality in the literature is mainly identified with Lee Carter model and its extensions.

As stated in [2] they could be framed into Generalized Age-Period-Cohort stochastic mortality models (GAPC).

# Generalized Age-Period-Cohort stochastic mortality models

In their definition they are close to widely used generalized linear models (GLM). As in GLM one needs to assume what distribution modelled risk follow.

For GACP number of deaths  $D_{xt}$  is treated as modelled phenomena.

Since number of deaths is an integer one could use:

- a) Binomial distribution (with logit link function)
- b) Poisson distribution (log link function)

# Stochastic mortality models

Linear formula that is then imputed into link function in its general form:

$$\eta_{xt} = \alpha_x + \sum_{i=1}^N B_x^{(i)} \kappa_t^{(i)} + B_x^{(0)} \gamma_{t-x}$$

- $\alpha_x$  stands for static age function capturing the general shape of mortality by age
- $N \geq 0$  is an integer indicating the number of age-period terms describing the mortality trends, with each time index  $\kappa_t^{(i)}$ ,  $i = 1, \dots, N$ , contributing in specifying the mortality trend and  $B_x^{(i)}$  modulating its effect across ages.
- The term  $\gamma_{t-x}$  accounts for the cohort effect with  $B_x^{(0)}$  modulating its effect across ages.

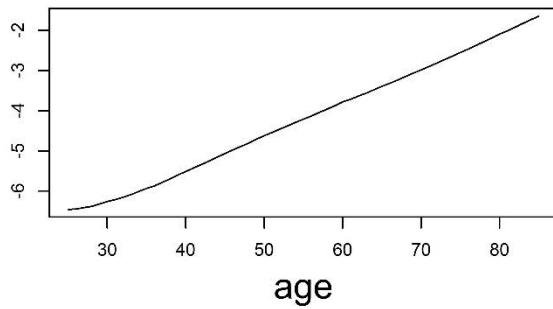
# Stochastic mortality models – Lee Carter

It is good to start with simplest and most popular model in this class defined under same notation. It could provide intuitive understanding of PLAT model for those already familiar with Lee-Carter.

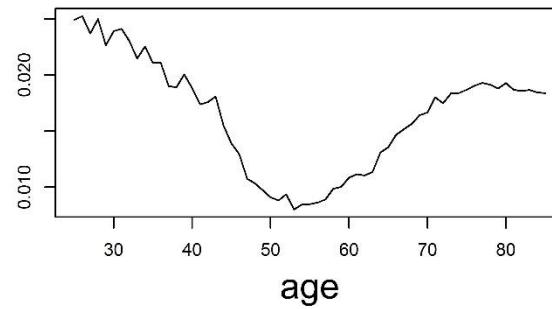
Lee-Carter model is defined as:

$$\eta_{xt} = \alpha_x + B_x^{(1)} \kappa_t^{(1)}$$

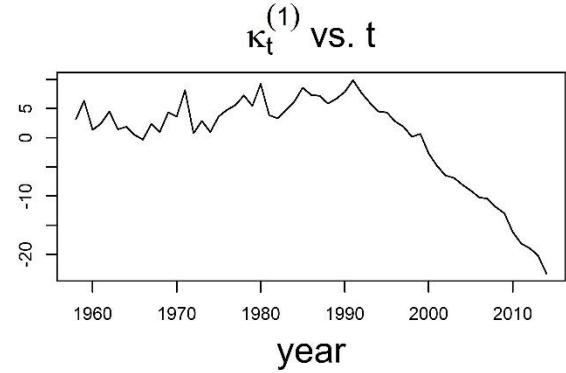
$\alpha_x$  vs.  $x$



$\beta_x^{(1)}$  vs.  $x$



$\kappa_t^{(1)}$  vs.  $t$

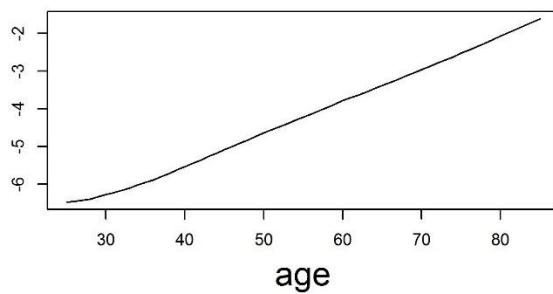


# Stochastic mortality models – PLAT

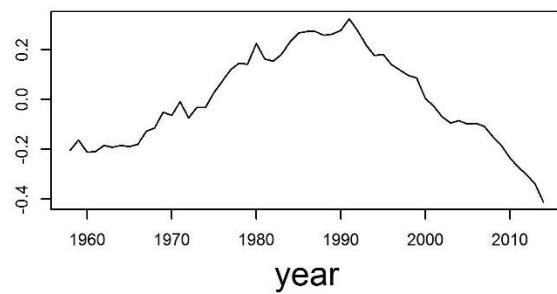
PLAT model is much more complex and defined as:

$$\eta_{xt} = \alpha_x + \kappa_t^{(1)} + (\bar{x} - x)\kappa_t^{(2)} + (\bar{x} - x)^+\kappa_t^{(3)} + \gamma_{t-x}$$

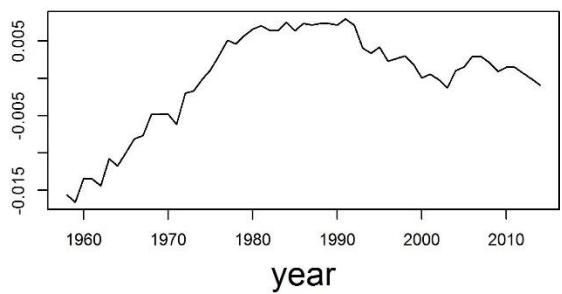
$\alpha_x$  vs. x



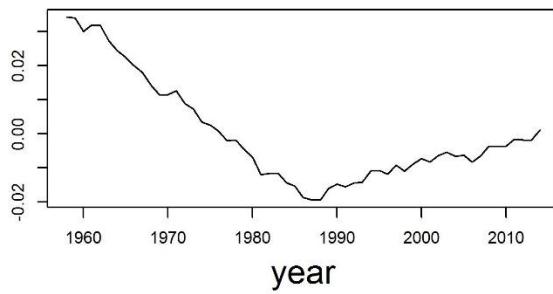
$\kappa_t^{(1)}$  vs. t



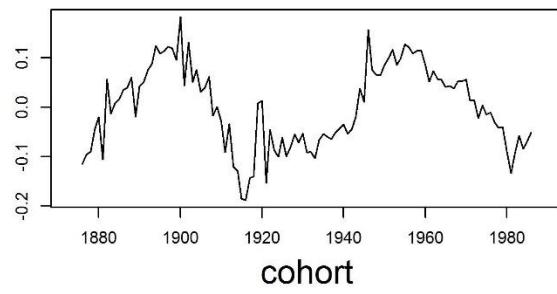
$\kappa_t^{(2)}$  vs. t



$\kappa_t^{(3)}$  vs. t



$\gamma_{t-x}$  vs. t-x



# Stochastic mortality models

In general steps that needs to be applied when working with those models are as follow:

- Fit model to data
- Assess goodness of fit
- Perform forecast (it could be perceives just as forecast of kappa and gamma parameters development over time)

# Forecast

Stochastic paths are generated by analytical transformation of fixed and forecasted parameters. Whole mortality projection could be reduced to the issue of forecast of kappa and gamma parameters.

It is assumed that number of deaths  $D_{xt}$  follow binomial distribution

$$D_{xt} \sim \text{Binomial}(E_{xt}^0, q_{xt})$$

Under that assumption  $q_{xt}$  could be calculated from analytical formula (logit link):

$$q_{xt} = \frac{e^{\eta_{xt}}}{(1 + e^{\eta_{xt}})}$$

Where  $\eta_{xt} = \alpha_x + \kappa_t^{(1)} + (\bar{x} - x)\kappa_t^{(2)} + (\bar{x} - x)^+\kappa_t^{(3)} + \gamma_{t-x}$  for PLAT model

Parameters that needs to be forecasted:  $\kappa_t^{(1)}, \kappa_t^{(2)}, \kappa_t^{(3)}, \gamma_{t-x}$

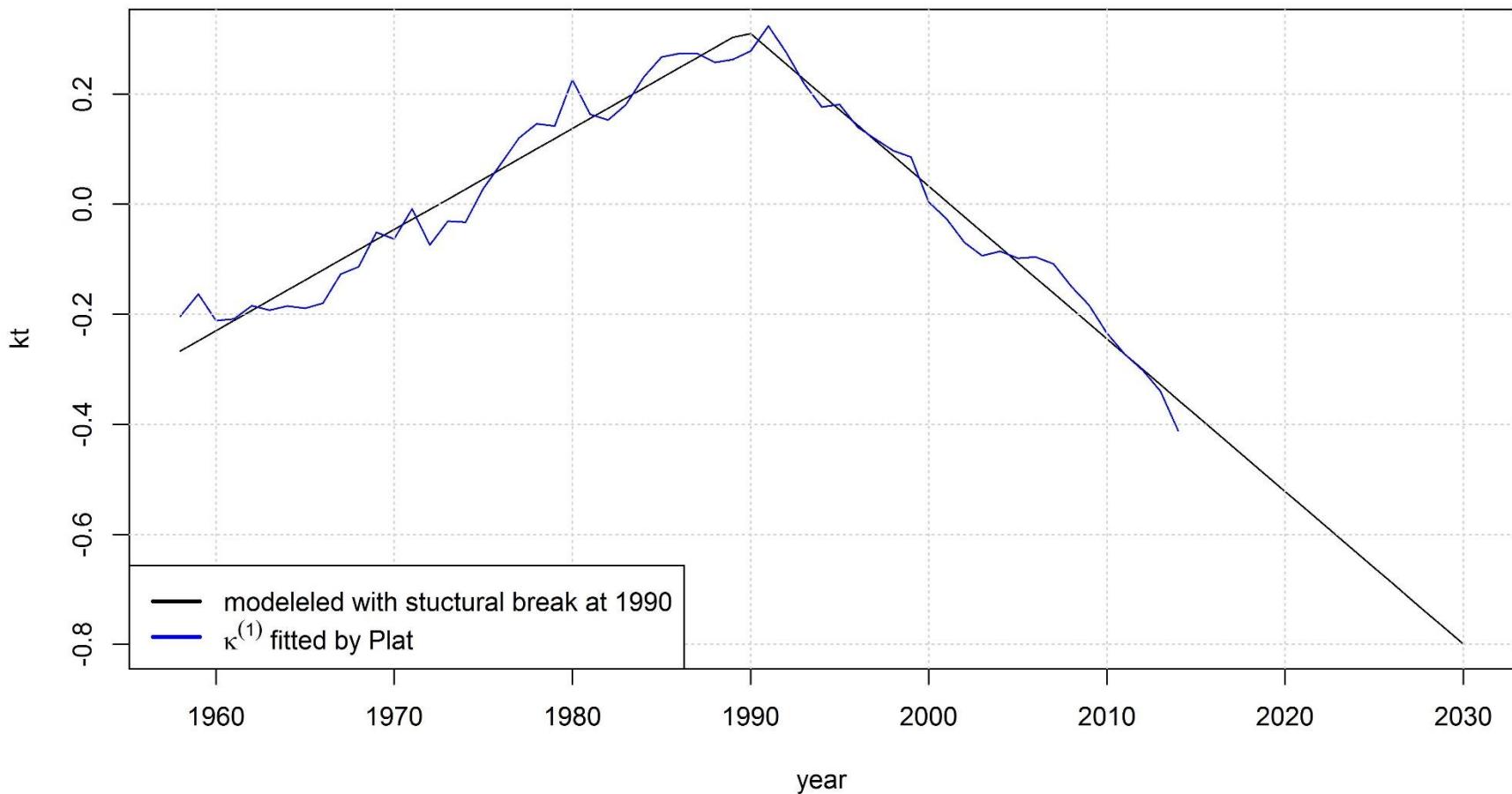
# Kappa parameter projection

Multivariate random walk with drift is applied (it is most popular approach for kappa parameters in literature).

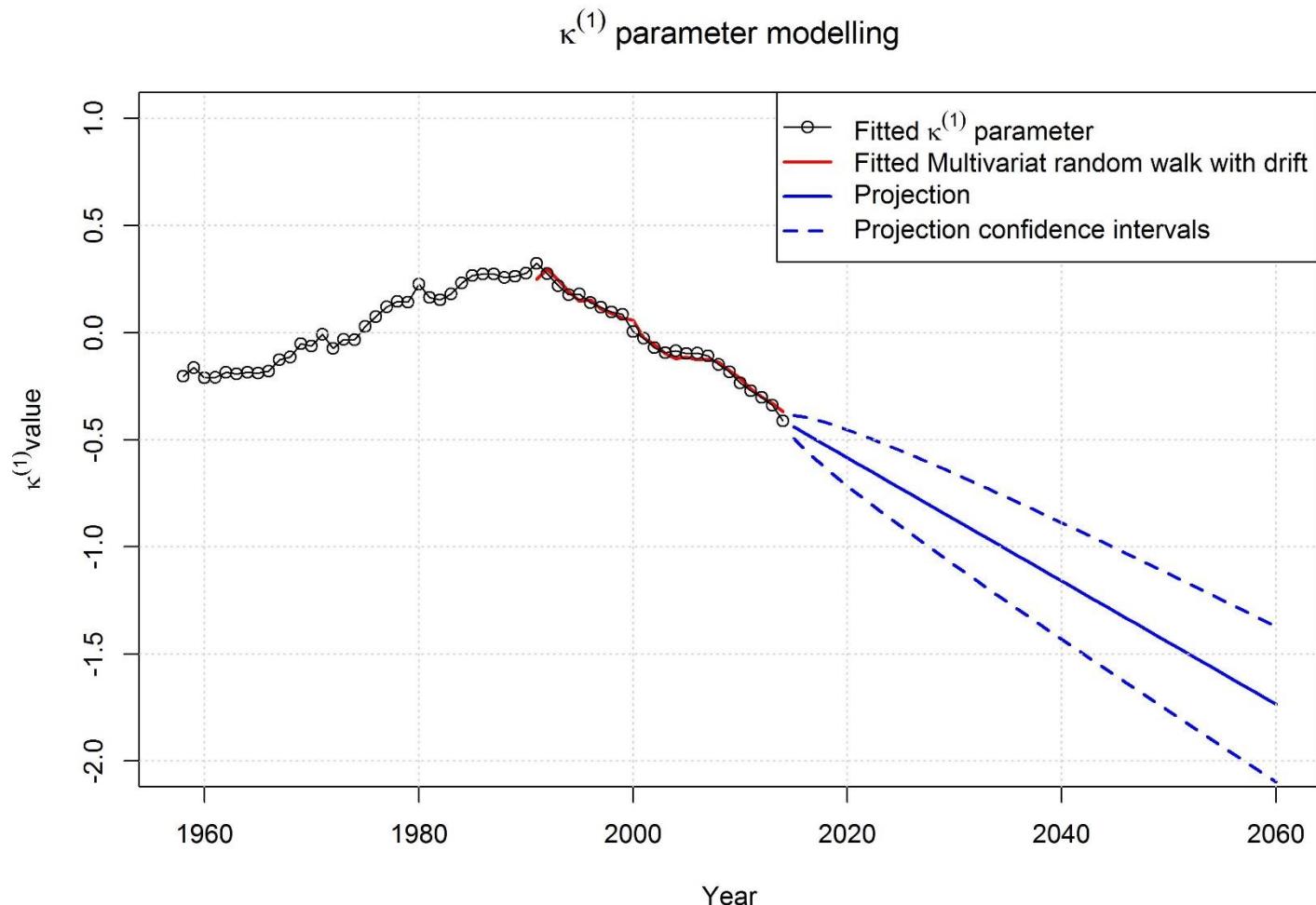
Also an structural break has been identified. Mathematical framework for performing structural change tests is provided within strucchange package.

# Structural break in 1990

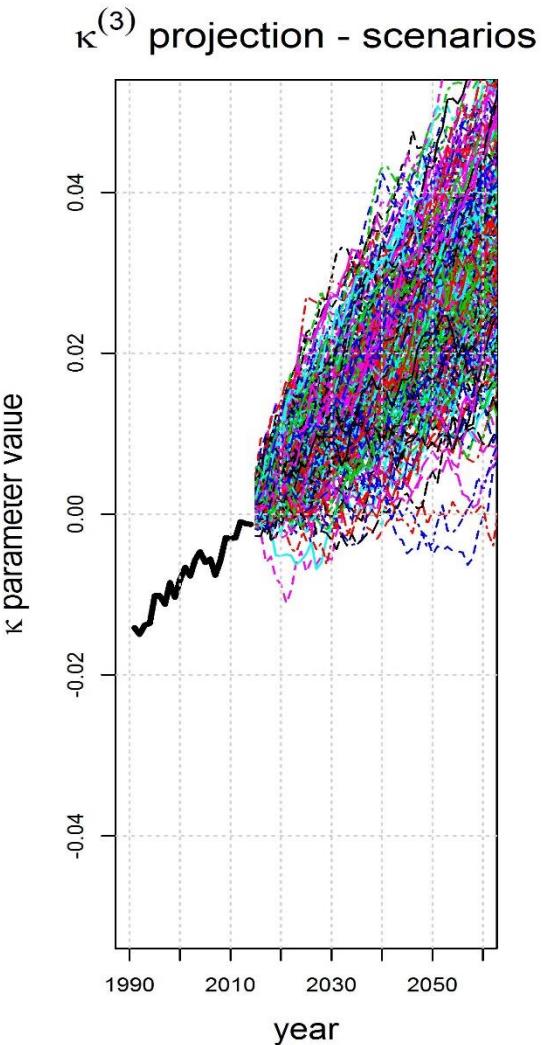
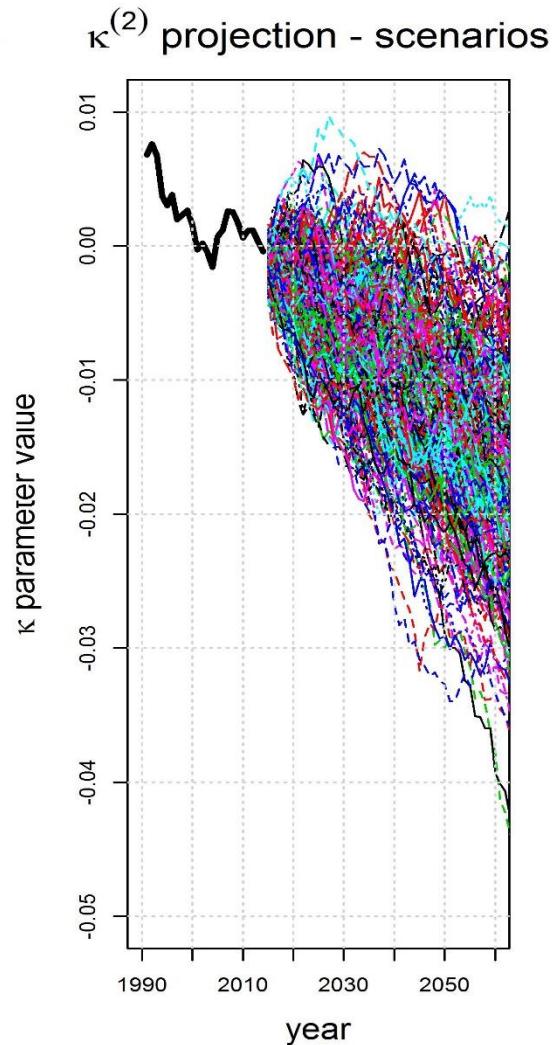
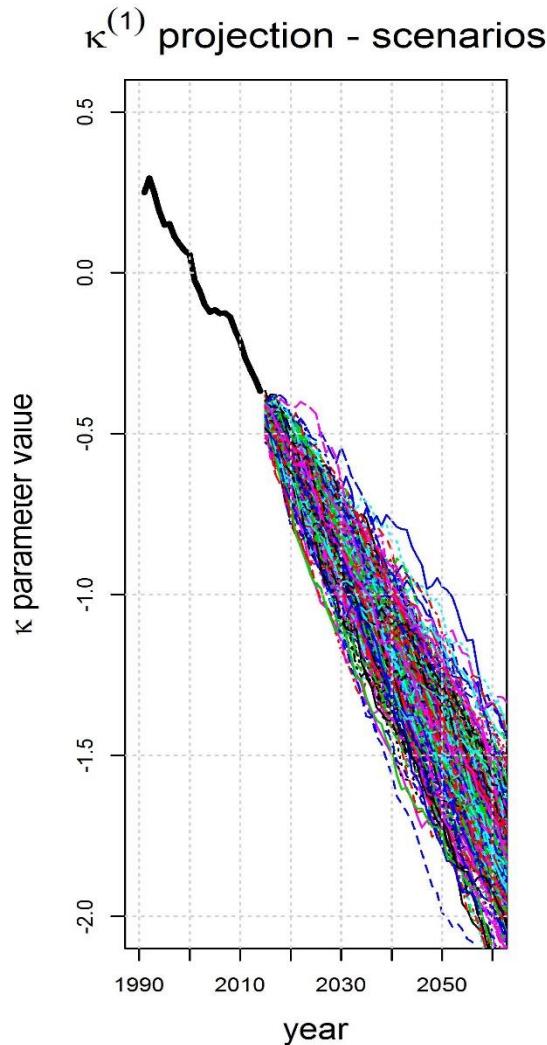
$\kappa^{(1)}$  parameter modeled with linear regression  $\kappa \sim t + SB + t^*SB$



# Structural break in 1990

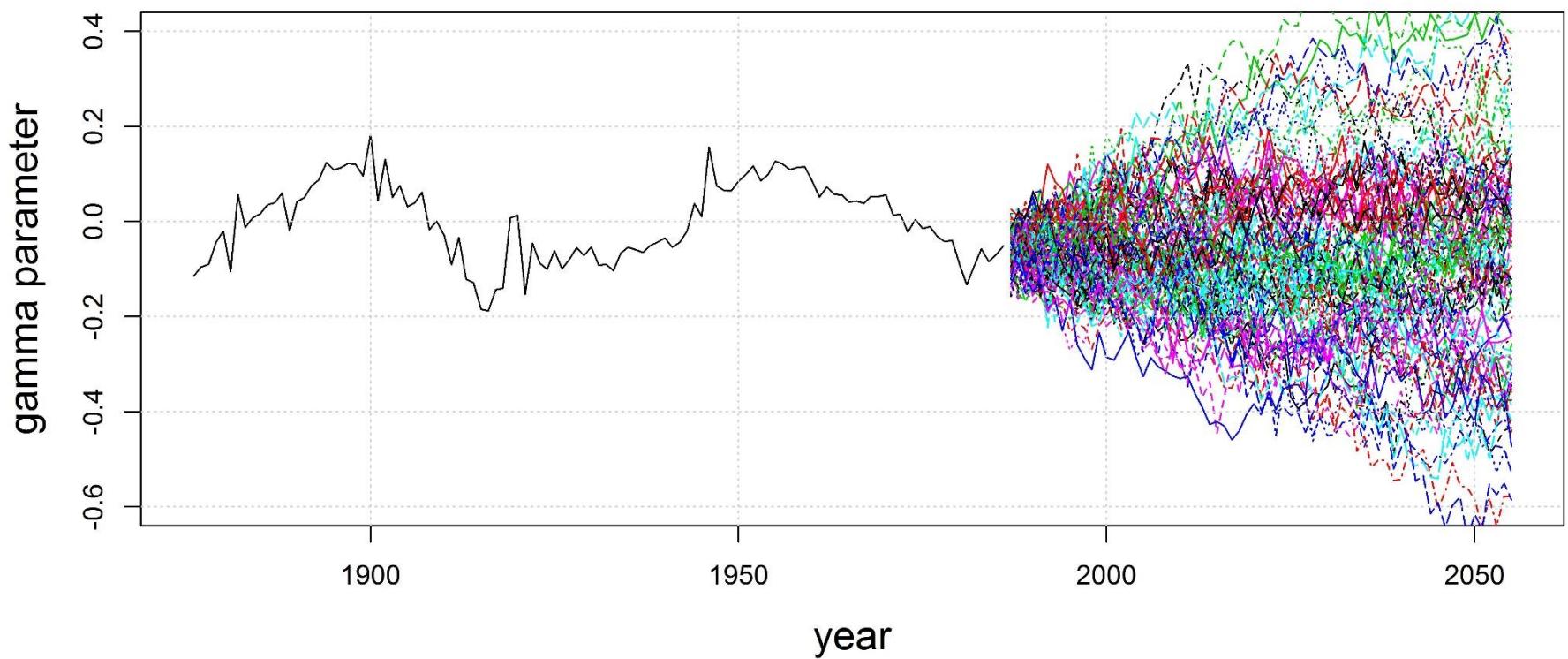


# Kappa parameters projection



# Gamma parameter projection

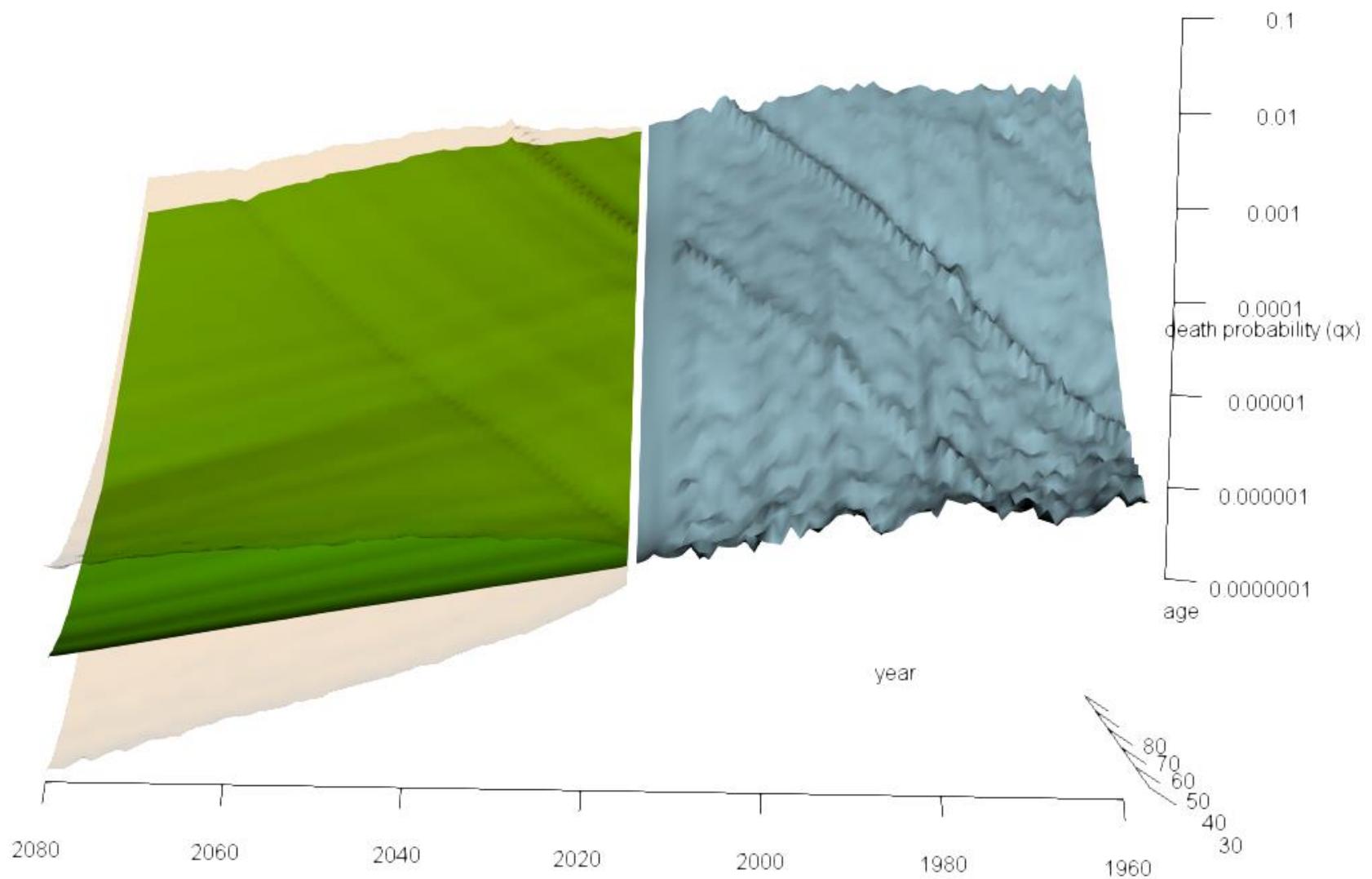
**gamma projection - scenarios**



# Projection – single age perspective



# Projection – whole surface perspective



# References

1. StMoMo: Stochastic Mortality Modelling, 2015  
<https://cran.r-project.org/web/packages/StMoMo/index.html>
2. Andrés M. Villegas, Vladimir Kaishev, Pietro Millossovich, 2015, StMoMo: An R Package for Stochastic Mortality Modelling  
<https://cran.r-project.org/web/packages/StMoMo/vignettes/StMoMoVignette.pdf>
3. Richard Plat, 2009, On Stochastic Mortality Modelling
4. Colin O'Hare, Youwei Li, 2014, Identifying Structural Breaks In Stochastic Mortality Models
5. Edviges Coelho, Luis C. Nunes, 2009, Forecasting Mortality in the Event of a Structural Change
6. Heather Booth, Leonie Tickle, 2008, Mortality modelling and forecasting: A review of methods
7. Zeileis A., Leisch F., Hornik K., Kleiber C., 2002, strucchange: An R Package for Testing for Structural Change in Linear Regression Models

# Contact data



wrobel.adam1990@gmail.com



<https://pl.linkedin.com/in/wrobeladam1>