

**MOODY'S**

**AXIS™ System Advanced Optimization Tools**

# Welcome

◆ **Moderator:** Martin Le Roux, Director – Actuary, Insurance

## Today's topic: Scenario Reduction with AXIS™ Machine Learning Tools

- Empower your application: Equipe attendees to confidently use AXIS™ scenario reduction tools
- Introduce the process combining human decisions, PCA, Clustering and Stratified Sampling for scenario reduction
- Demonstrate the method's accuracy and superiority to random sampling using case study results.

## Meet the Speakers

Yulia Nezhlin: 14 years at Moody's, expert in ML for actuarial use, PhD in Physics

## ☀️ Why we're here

- **Modernizing the AXIS™ system:** We're continuously enhancing the AXIS™ Actuarial System to help you work faster and smarter with large, complex models.
- **Empowering actuaries:** Our built-in tools are evolving to support efficient, high-impact analysis—driving better outcomes with less effort.

## Ask us your questions

Submit questions via the Q&A box

**Save the date: Insurance Solutions user group**

Sunday, August 17, 2025 | 2:00 PM - 5:00 PM CT at Hotel EMC2, Autograph Collection, Chicago, IL

Ahead of the SOA ValAct conference

**MOODY'S**

Yulia Nezlin

**Scenario Reduction with AXIS™ Machine Learning Tools**

# Contents

## 1. Scenario Reduction Methodology in AXIS™ System

- Human Decisions
- Automative Processing

## 2. Demonstrations

- Scenario Reduction Preserving Risk-neutrality
  - See KB 2548
  - Sample Dataset
- GMWB Case Study
  - KB 2560

# Objective, Recent Achievements, New approach

## Run-time problems in Monte Carlo simulations

- Too many economic scenarios needed to satisfy accuracy requirements

**Objective:** Reduce the number of scenarios while preserving the multidimensional distribution of target functions of interest (e.g. PV of projected cashflows)

- **AAA Scenario Picking Tool**

- developed for interest rate scenarios
- A qualitative approach to define a single significant feature (“significance”)
- Based on stratification by one feature (significance)

- **The new (machine learning) approach**

- generalizes the AAA approach for a wider class of applications (not limited to interest rate scenarios and not limited to applications with a single significant feature)
- Extracts significant features from scenario data using machine learning tools
- Based on multidimensional stratification by ‘a number of’ significant features

# The Method

- **Human Decisions**

Control automatic processing

- Prepare a proper input to **Machine Learning**
- Make PCA applicable
- Might require a modeling (approximations)

- **Machine Learning**

- **PCA**

(Principal Component Analysis)

- Reduces dimension
  - Efficient because of correlations within scenario data
  - Makes clustering applicable
- Limitation: works only on LINEAR data

- **Clustering**

- Groups together ‘similar’ scenarios
- Limitation: high dimension

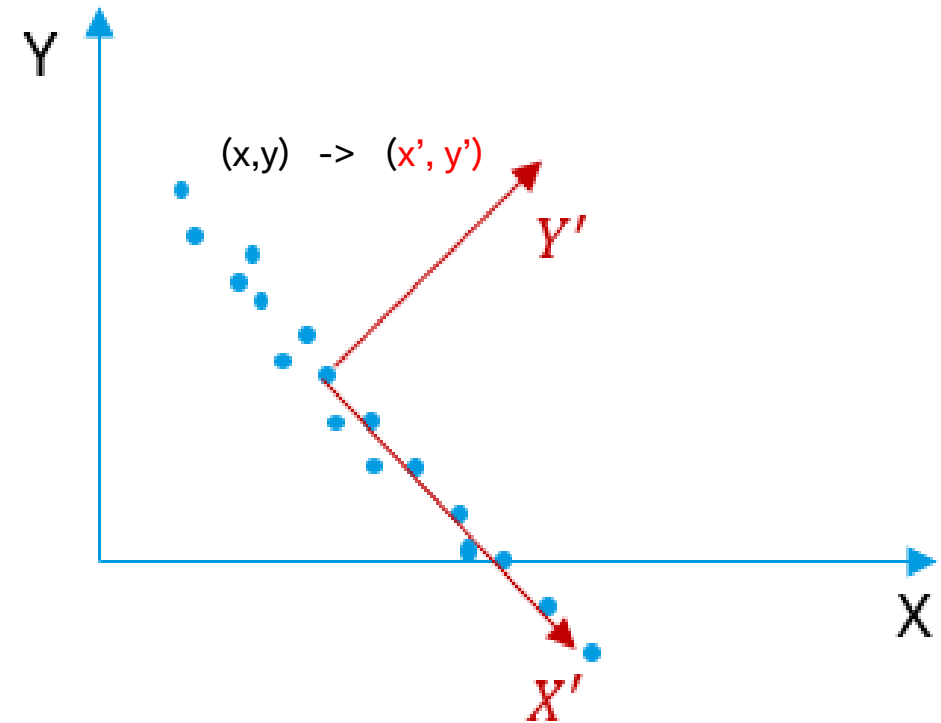
- **Stratified Sampling**

- Reduces sampling error compared to random selection
- Requires pre-stratification (clustering)

# Supplementary Information

## PCA Illustration

Two-dimensional data points  $(x,y)$



$x$  and  $y$  are correlated

Perform a linear transform:  $(X,Y) \rightarrow (X',Y')$

$x'$  and  $y'$  are uncorrelated (*Principal Components*)

$x'$  has much bigger variance – First Principal Component (PC1)  
 $y'$  has smaller variance – Second Principal Component (PC2)

PC1 – is the most significant feature of those datapoints  $(x,y)$

# A simple Demonstration -

Sampling to preserve the risk  
neutrality for a wide range of  
time projections

see also KB 2548

## Given 5,000 Scenarios:

- Interest rate market, 2 Equity markets
- 30 years annual

## Objective:

Pick a subset of 500 scenarios out of 5000 to preserve the distribution of

- $Martingale1(t, scn) = Eq\_Index1(t, scn) / risk\_free\_BA(t, scn), \quad t = 1, \dots, 30$ 
  - $\langle Martingale1(t, scn) \rangle = 1$
- $Martingale2(t, scn) = Eq\_Index2(t, scn) / risk\_free\_BA(t, scn), \quad t = 1, \dots, 30$ 
  - $\langle Martingale2(t, scn) \rangle = 1$

**Target Function**  $(scn) = (Martingale1(t, scn), Martingale2(t, scn)), \quad t = 1, \dots, 30$   
- 60 components

## Why this is simple?

Because the connection between the Target Function and the raw scenario components is known exactly.

## Why this is hard?

Because the dimension of the target function is high (60).

# A simple Demonstration -

Sampling to preserve the risk  
neutrality for a wide range of  
time projections

Macro Batch in AXIS™ System

The screenshot displays the 'Dataset [ CLUSTERING SCENARIOS 2 MARTINGALES ] - Batch' window. It contains a table with the following data:

Id	* Batch Name[+]	Time Stamp	Run Time Stamp
25	1. Human Decision: create 60 Martingale variables	27/05/2025 04:52 PM	20/06/2025 10:12 PM
27	2. Run PCA	28/05/2025 02:48 PM	28/05/2025 02:48 PM
57	3. Run Clustering	22/05/2025 02:31 PM	16/06/2025 12:15 PM
37	4. Run Stratified Sampling	28/05/2025 02:47 PM	17/06/2025 05:34 PM
30	DEMO scenario sampling macro	22/06/2025 04:58 PM	27/05/2025 07:00 PM

Below the table is the 'Dataset Macro Wizard' dialog box. It shows 'Batch processes selected for Dataset Macro:' with a table of selected items:

Reference	Id	Name	Enabled
0	25	1. Human Decision: create 60 Martingale variables	<input checked="" type="checkbox"/>
1	27	2. Run PCA	<input checked="" type="checkbox"/>
2	57	3. Run Clustering	<input checked="" type="checkbox"/>
3	37	4. Run Stratified Sampling	<input checked="" type="checkbox"/>

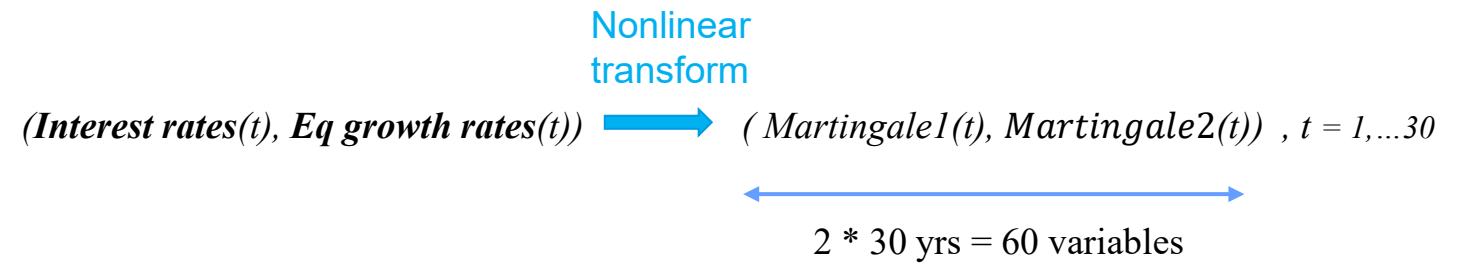
The wizard also includes a checkbox for 'Use the same Run Id for all batch jobs' (unchecked) and several control buttons: Add..., Insert..., Remove, Up, Down, Edit..., Enable All, and Disable All. At the bottom are Help, Cancel, < Back, Next >, and Finish buttons.

# A simple Demonstration –

sampling to preserve the risk  
neutrality for a wide range of  
time projections

1. Batch Human Decisions:  
Create Martingale variables

## 1. Transform raw scenario features to new variables (suitable for PCA):

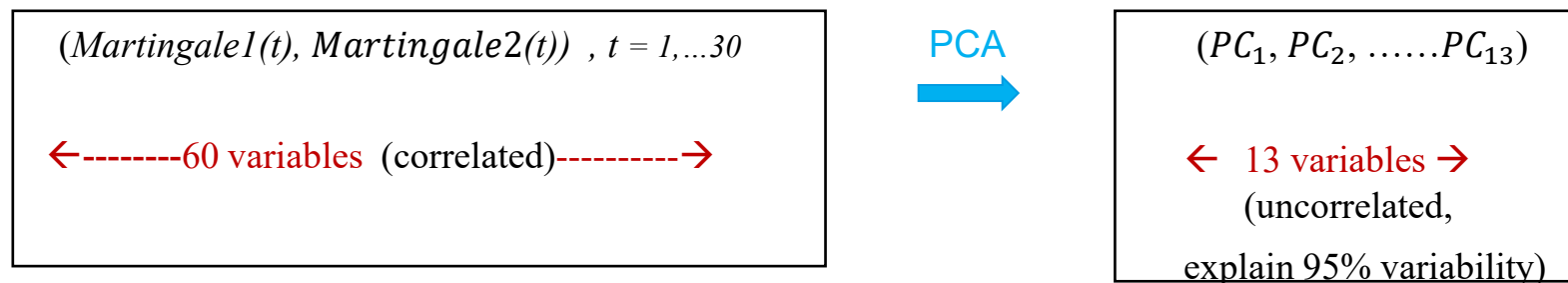


# A simple Demonstration –

sampling to preserve the risk  
neutrality for a wide range of  
time projections

## 2. Batch Run PCA

**2. Perform PCA on 60 new variables and leave only several Principal Components that preserve the most of variability:**



PCA replaces numerous highly correlated variables (Martingale ratios) with a much smaller number of uncorrelated ones (Principal Components)

**The most significant features - 13 Principal Components**

# A simple Demonstration –

sampling to preserve the risk  
neutrality for a wide range of  
time projections

## 3. Batch Run Clustering

### **Cluster scenarios by 13 Principal Components (Location Variables)**

- Create 50 clusters (about 100 scenarios per cluster)
  - Why 50?
    - Estimate: 500 subset scenarios / 10 scenarios per cluster = 50

# A simple Demonstration –

sampling to preserve the risk  
neutrality for a wide range of  
time projections

## 4. Batch Run Stratified Sampling

### **Use Stratified Sampling to pick 500 scenarios from 50 clusters:**

- Pick scenarios independently and randomly from each cluster
- The number of representative scenarios from each cluster is proportional to the size of the cluster

# A simple Demonstration –

sampling to preserve the risk  
neutrality for a wide range of  
time projections

## 5. Testing the results

The result is expected to be more accurate than random pick

### To test:

- Compute the distributions of  $Martingale1(t,scn)$ ,  $Martingale2(t,scn)$  over 500 scenarios and compare with the benchmark distribution from 5000 scenarios
  - Compute the errors
- Solution is not unique
  - Repeat macro-batch 100 times with random seed (in K-means clustering and Stratified Sampling)
  - Compute MSE (mean square error)
- Plot the errors

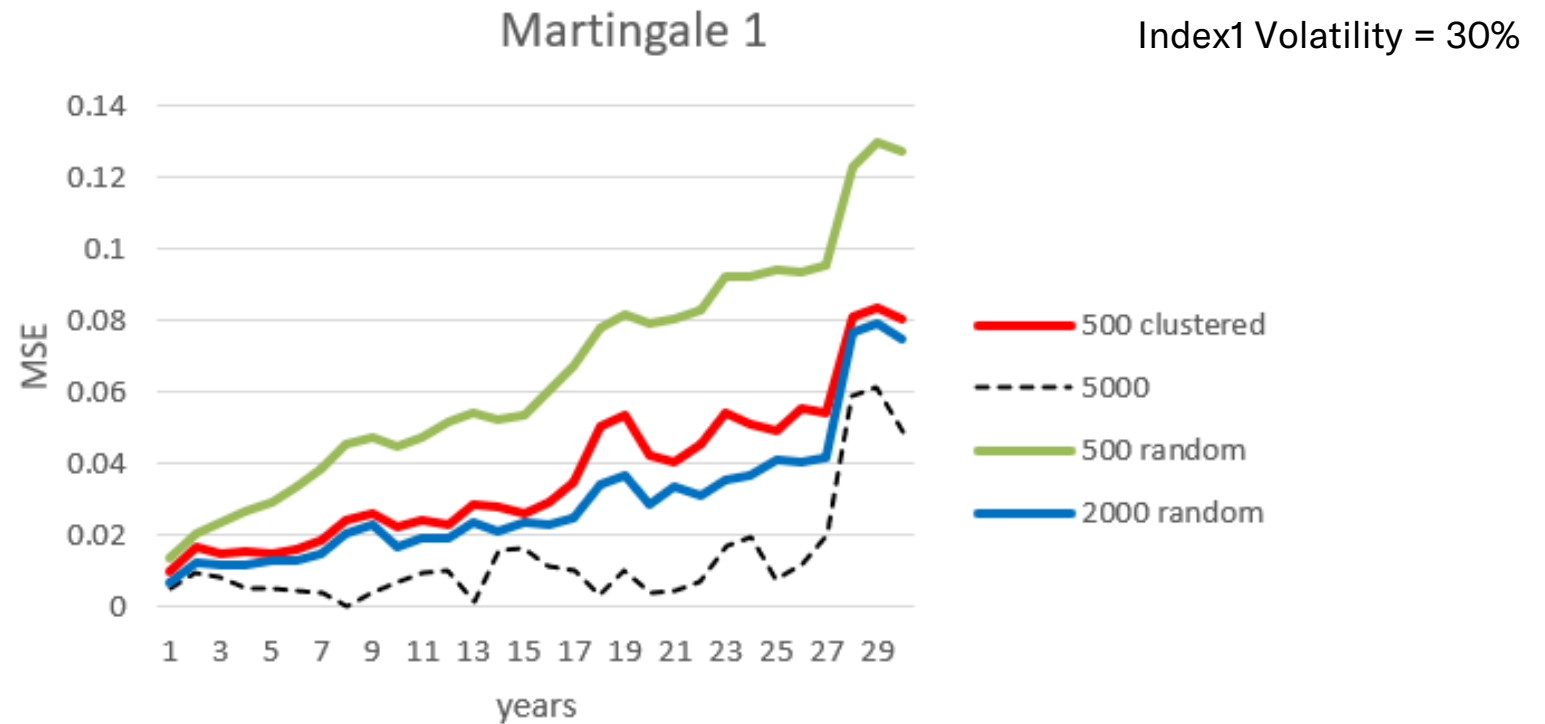
# A simple Demonstration – testing marginal distributions of Martingale1

Mean Square Errors (MSE) of marginal means from 100 samples

- The MSEs of the marginal means from 500 selected scenarios are smaller than those from 500 random scenarios
- 500 selected scenarios are 'equivalent' to ~ 2000 random scenarios
- Compression: ~4 times

Efficiency depends on correlations within data

- Defined by scenario generator model

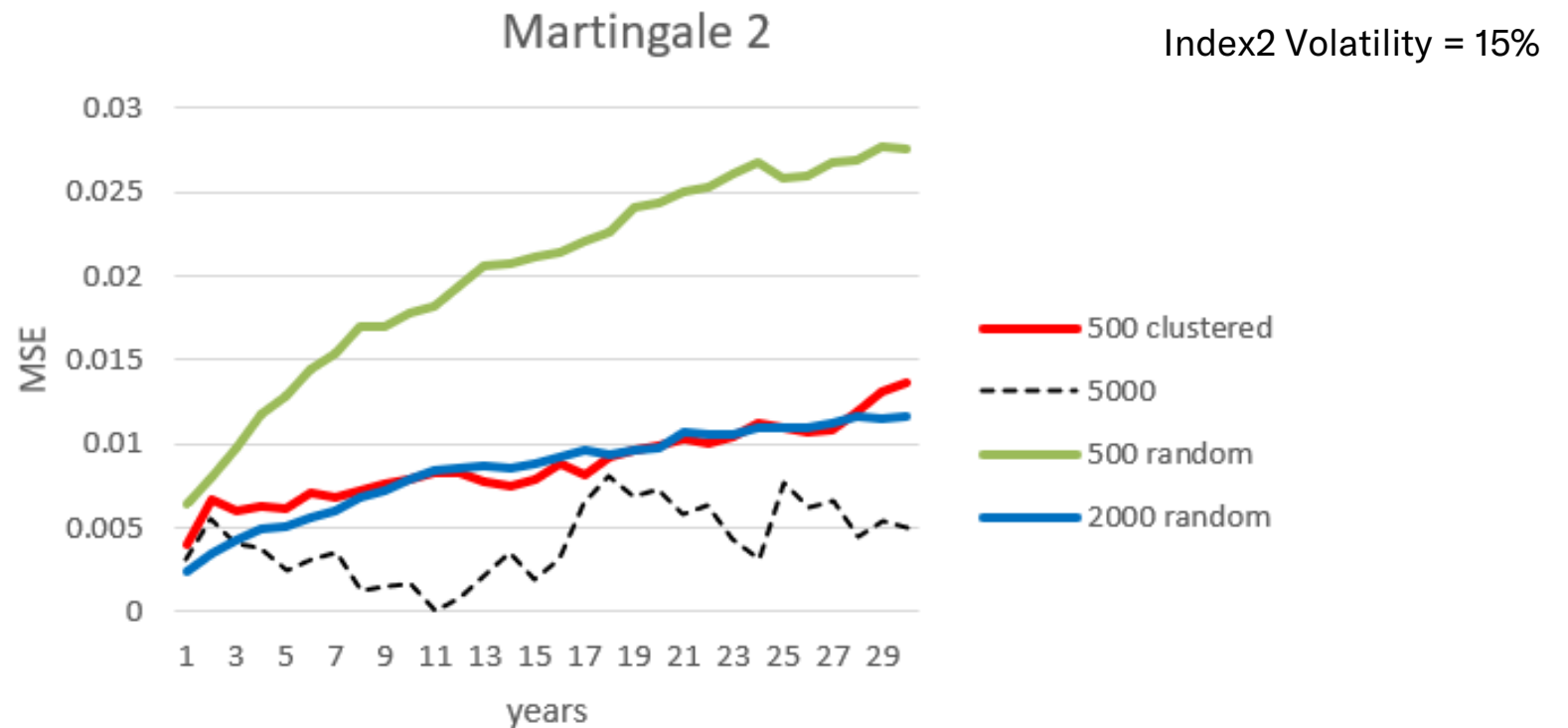


# A simple Demonstration– testing marginal distributions of Martingale2

Mean Square Errors (MSE) of marginal means from 100 samples

- The MSEs of the marginal means from 500 selected scenarios are smaller than those from 500 random scenarios
- 500 selected scenarios are 'equivalent' to about 2000 random scenarios
- Compression: ~4 times

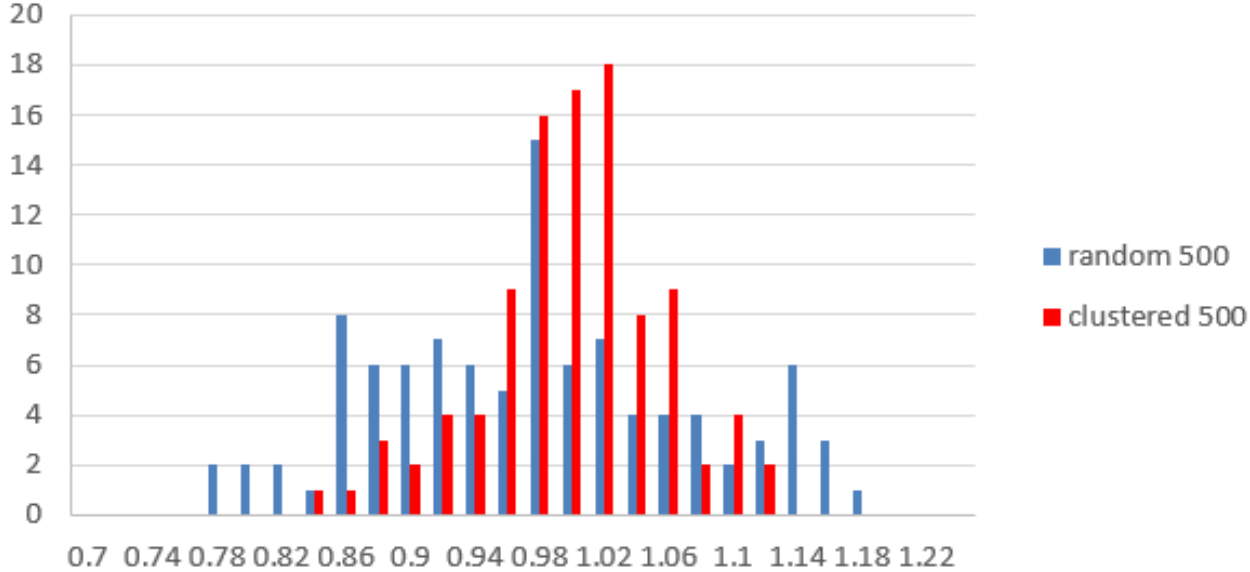
- Efficiency depends on correlations within data
- Defined by scenario generator model



# A simple Demonstration – testing marginal distributions of Martingale ratio

- More narrow distribution of the mean of 500 clustered scenarios compared to 500 random scenarios

Year 30 Marginal Mean Distribution of Martingale 1 over 100 simulations

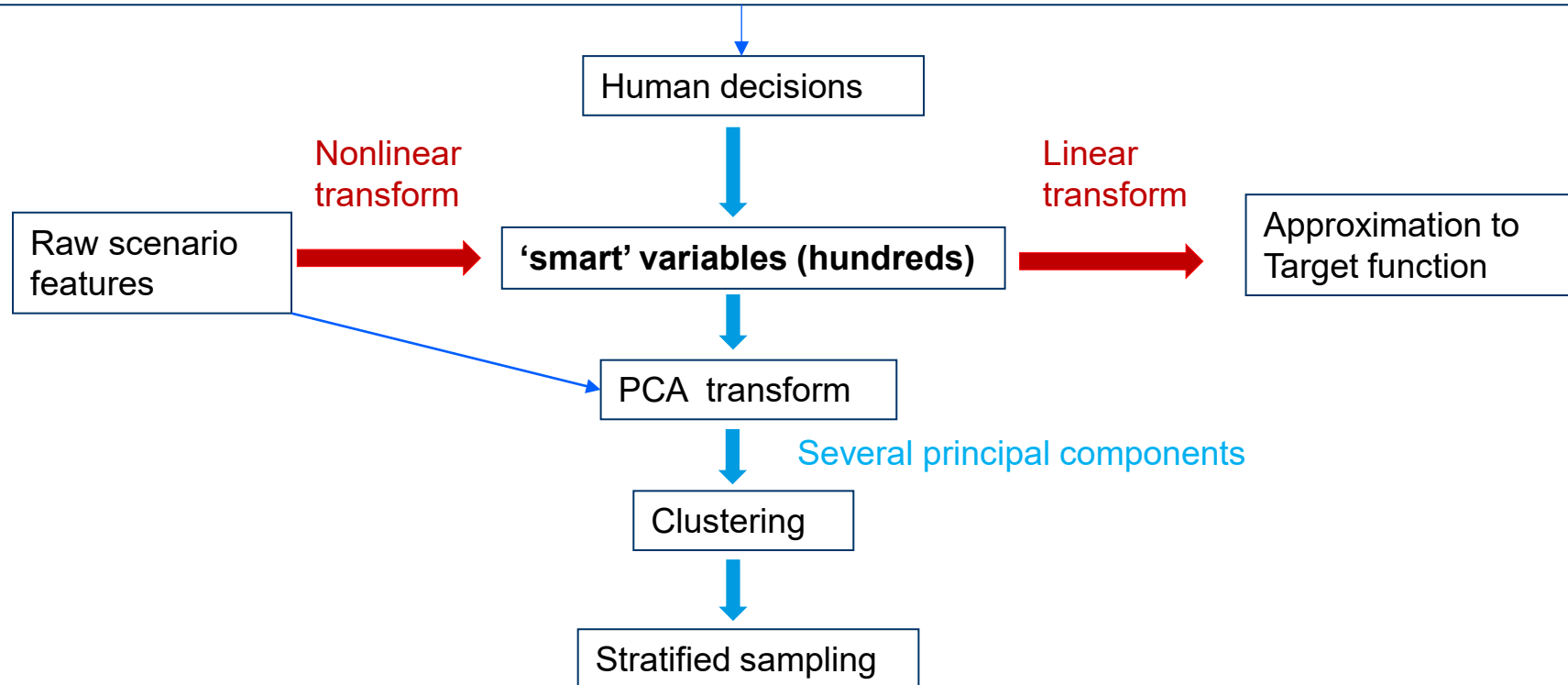


# What to do in a more general case?

$$\text{Target CF}(\text{scenario}) = \sum_{scen} \sum_{pol}^n F(\text{Pol}, \text{Scenario}),$$

## Human decisions:

- Create new ('**smart**') variables (can be hundreds), a linear function of which will approximate the target function
- Might require some modeling



# Straightforward Human Decisions: Smart Variables Examples

- PV of fixed Cash Flows (AAA picking tool applications):

$$Target(Scenario) = \sum_{pol=1}^N \sum_{mon=1}^{120} CF(pol, mon) * \boxed{Discount(mon, Scenario)}$$

smart variables

The most significant features are determined by Principal Components ( $Discount(mon, Scenarios)$ )

- PV of equity dependent Cash Flows,  $CF(mon, Index(Scenario))$ :

$$Target(Scen) = \sum_{pol=1}^N \sum_{mon=1}^{120} F(pol, mon) * \boxed{Index(mon, Scen) * Discount(mon, Scen)}$$

smart variables

- These smart variables are correlated and thus, dimension can be reduced by PCA

# Case Study: GMWB Simplified model

See also KB 2560

- 128 policies
- The GMWB coverages are linked to an equally weighted set of five equity indices
- 3 years ratchet guarantee
- 10,000 scenarios containing 60 years of monthly data
- A subset of 500 scenarios was selected using the described methodology
  - 5 Principal Components explain 95% variability
  - 50 clusters

# Case Study: GMWB Simplified model

See also KB 2560

- **Target function is the PV of GMWB Benefits**

$PV\_Benefits(Scn, Policy) \sim$

$$\sum_{m=T(Scn)}^{720} Guarantee(T, Scn, Policy) * Dscnt(m, Scn) * Lx(m, Policy),$$

- **Approximation:**

$Guarantee(T, Scn, Pol) = G(Fund(T, Scn, Pol), Fund(T-1, Scn, Pol), \dots, Fund(1, Scn, Pol))$

$\sim Linear(Fund(m, Scn, Policy)), \quad m=1, \dots, T$

$Fund(m, Scn, Policy) = Linear(Index(m, Scn))$

Model  
parameters

$PV\_Benefits(Scn, Policy) \approx Linear[Index(m, Scn) * Dscnt(m + lag, Scn)], \quad m=1, \dots, M$

In our model :  $M= 180, lag = 0$

- **Results:**

For the mean

Standard Error with 500 stratified scenarios	Standard Error with 500 random scenarios	Standard Error with Original 10,000 random scenarios
0.48%	1.94%	0.43%

# Summary

## Key takeaways

**AXIS™ offers Optimization Tools to efficiently process and compress scenario data:**

- **Automated scenario reduction uses PCA, Clustering, and Stratified Sampling techniques.**
- **Human interaction is essential for effective scenario reduction.**

## Next steps

- Please reach out to your Moody's contact if you would like to discuss any of this further.
- Don't forget to join us at our user group in Chicago, ahead of the SOA ValAct conference.
- Recording, slides, and Q&A will be shared within 48 hours.
- Complete the survey to share your feedback.

**Save the date: Insurance Solutions user group**  
Sunday, August 17, 2025 | 2:00 PM - 5:00 PM CT at  
Hotel EMC2, Autograph Collection, Chicago, IL  
Ahead of the SOA ValAct conference



**Scenario Reduction  
with  
AXIS™ Machine  
Learning Tools**

**Thank you**