# Dependence modeling in General Insurance using LGC and HMMs

Zabibu Afazali
(Makerere University, Uganda)

July 26, 2023

PhD Forum Presentation

Actuarial, Finance, Risk, and Insurance Congress(AFRIC)

# Dependence modeling-Motivation

- Diversification, reserving,pricing, reinsurance

# Dependence modeling-Motivation

- Diversification, reserving,pricing, reinsurance
- Non-linear and tail dependence- an issue?

# Why LGC and HMMs?

**LGC**

- Linear to non-linear $\rho$ interpretation

**HMMs**

# Why LGC and HMMs?

**LGC**

- Linear to non-linear $\rho$ interpretation
- Comprehensible visualizations

**HMMs**

# Why LGC and HMMs?

**LGC**

- Linear to non-linear $\rho$ interpretation
- Comprehensible visualizations
- Entire dependence structure by set of pairwise correlations

**HMMs**

# Why LGC and HMMs?

**LGC**

- Linear to non-linear $\rho$ interpretation
- Comprehensible visualizations
- Entire dependence structure by set of pairwise correlations
- Guassian densities approximated locally using a bivariate correlation function

**HMMs**

# Why LGC and HMMs?

**LGC**

- Linear to non-linear $\rho$ interpretation
- Comprehensible visualizations
- Entire dependence structure by set of pairwise correlations
- Guassian densities approximated locally using a bivariate correlation function

**HMMs**

- Capture complex temporal dependencies

# Why LGC and HMMs?

**LGC**

- 

# Why LGC and HMMs?

## LGC

- Linear to non-linear $\rho$ interpretation
- Comprehensible visualizations
- Entire dependence structure by set of pairwise correlations
- Guassian densities approximated locally using a bivariate correlation function

## HMMs

- Capture complex temporal dependencies
- Handling Varying Patterns and Regimes.
- Assumes that data-generating process corresponds to a time-dependent **mixture** of conditional distributions driven by **Hidden states**

# LGC basic notation

**Define**

- Let $X = (X_1, X_2)$ 2D RV with density $f(x) = f(x_1, x_2)$ and $\psi$ be a Guassian bivariate
- $\mu(x) = (\mu_1(x), \mu_2(x))$ is the local mean vector
- $\Sigma(x) = \sigma_{ij}(x)$ is the local covariance matrix.

# LGC basic notation

**Define**

- Let $X = (X_1, X_2)$ 2D RV with density $f(x) = f(x_1, x_2)$ and $\psi$ be a Guassian bivariate
- $\mu(x) = (\mu_1(x), \mu_2(x))$ is the local mean vector
- $\Sigma(x) = \sigma_{ij}(x)$ is the local covariance matrix.
- $\rho(x)$ is the correlation coe□cient in $\psi$

# LGC basic notation

**Define**

- Let $X = (X_1, X_2)$ 2D RV with density $f(x) = f(x_1, x_2)$ and $\psi$ be a Guassian bivariate
- $\mu(x) = (\mu_1(x), \mu_2(x))$ is the local mean vector
- $\Sigma(x) = \sigma_{ij}(x)$ is the local covariance matrix.
- $\rho(x)$ is the correlation coefficient in $\psi$
- Approximate $f$ locally at each point $x$ by $\psi_x(v)$
- $v = (v_1, v_2)^T$ is the running variable
- The local population parameters $\lambda(x) = (\mu(x), \Sigma(x))$ can be defined by minimizing a penalty function $q$ measuring the difference between $f$ and $\psi$

# HMMs basic notation

- $C_t : t = 1, \ldots, T$ hidden states,

# HMMs basic notation

- $C_t : t = 1, \ldots, T$ hidden states,
- $R_t : t = 1, \ldots, T$ observed variables

# HMMs basic notation

- $C_t : t = 1, \ldots, T$ hidden states,
- $R_t : t = 1, \ldots, T$ observed variables
- $\Gamma = \gamma_{ij}$ TPM and $\delta$ stationary distribution

# HMMs basic notation

- $C_t : t = 1, \ldots, T$ hidden states,
- $R_t : t = 1, \ldots, T$ observed variables
- $\Gamma = \gamma_{ij}$ TPM and $\delta$ stationary distribution
- $R^{(t)}$ the “history” of the observed process

# HMMs basic notation

- $C_t : t = 1, \ldots, T$ hidden states,
- $R_t : t = 1, \ldots, T$ observed variables
- $\Gamma = \gamma_{ij}$ TPM and $\delta$ stationary distribution
- $R^{(t)}$ the “history” of the observed process
- $\theta$ vector of model parameters

# HMMs basic notation

- $C_t : t = 1, \ldots, T$ hidden states,
- $R_t : t = 1, \ldots, T$ observed variables
- $\Gamma = \gamma_{ij}$ TPM and $\delta$ stationary distribution
- $R^{(t)}$ the “history” of the observed process
- $\theta$ vector of model parameters
- Express $L(\theta)$ likelihood of observations

# HMMs basic notation

- $C_t : t = 1, \dots, T$ hidden states,
- $R_t : t = 1, \dots, T$ observed variables
- $\Gamma = \gamma_{ij}$ TPM and $\delta$ stationary distribution
- $R^{(t)}$ the “history” of the observed process
- $\theta$ vector of model parameters
- Express $L(\theta)$ likelihood of observations
- Compute $L(\theta)$ using forward (or Backward) algorithm

# HMMs basic notation

- $C_t : t = 1, \ldots, T$ hidden states,
- $R_t : t = 1, \ldots, T$ observed variables
- $\Gamma = \gamma_{ij}$ TPM and $\delta$ stationary distribution
- $R^{(t)}$ the “history” of the observed process
- $\theta$ vector of model parameters
- Express $L(\theta)$ likelihood of observations
- Compute $L(\theta)$ using forward (or Backward) algorithm
- Used TMB


QUESTION?

# The Data

- Two data sets from Kenya and Norway, monthly and then weekly

# The Data

-

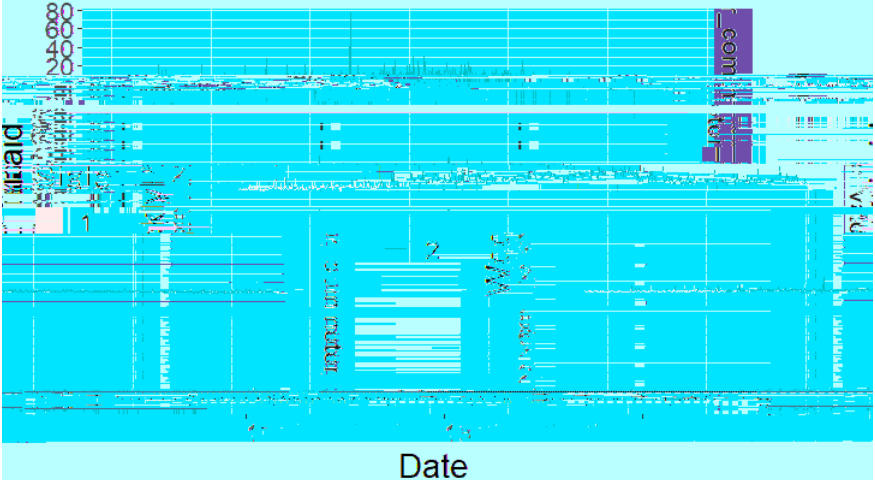
# The Data

- Two data sets from Kenya and Norway, monthly and then weekly
- 7 lines of business considered(Personal Accident, Fire industrial, Motor private, motor comprehensive, Workers compensation,liability, Engineering)
- LGC and HMM Illustrated using Motor LoBs and Homeowners Insurance LoBs

# The Data

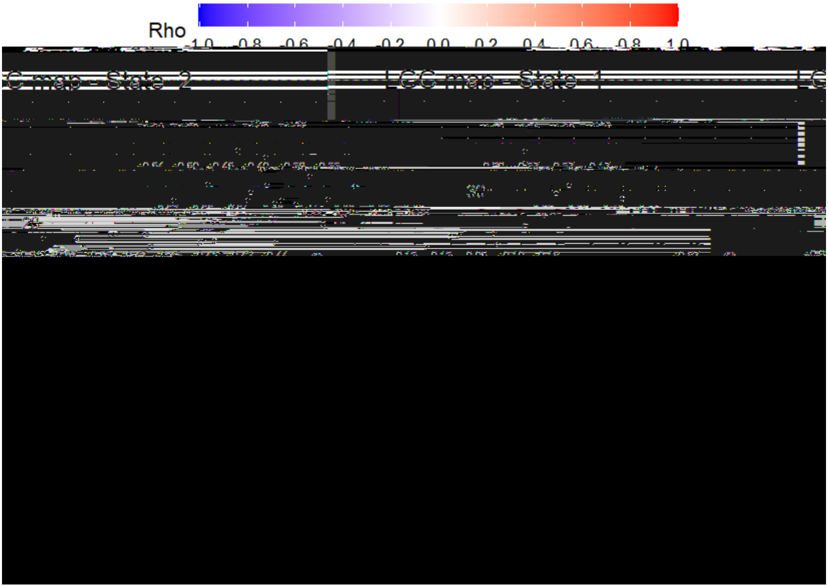
- Two data sets from Kenya and Norway, monthly and then weekly
- 7 lines of business considered(Personal Accident, Fire industrial, Motor private, motor comprehensive, Workers compensation,liability, Engineering)
- LGC and HMM Illustrated using Motor LoBs and Homeowners Insurance LoBs
- For motor-15 years from July 2007 to Dec 2021 giving 756 weekly records.

motor_commercial vs motor_private, rho = 0.85

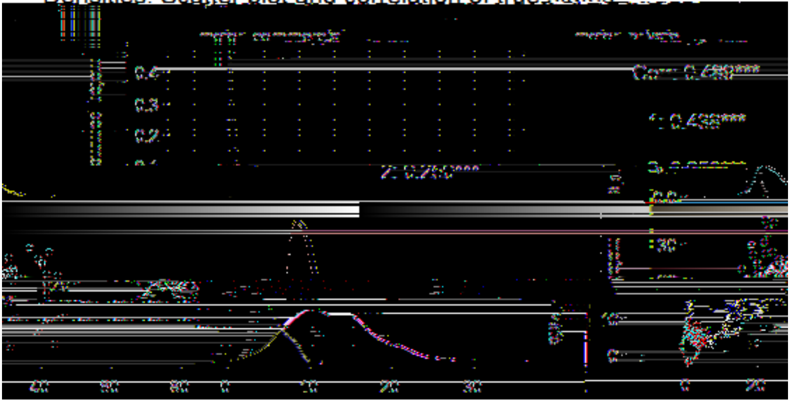

80
60
40
20
Date

## Table: Table central measures

| Descriptive measures | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| LoB | States | n | mean | sd | median | min | max | Skew | Kurtosis | se |
| Motor-commercial | state 1 | 615 | 4.41 | 3.53 | 3.74 | -1.23 | 15.789 | 0.69 | -0.33 | 0.14 |
| | state 2 | 141 | 11.65 | 9.14 | 10.53 | -12.59 | 77.03 | 2.62 | 16.79 | 0.77 |
| | All states | 756 | 5.76 | 5.79 | 4.44 | -12.59 | 77.03 | 3.42 | 30.96 | 0.21 |
| Motor-private | state 1 | 615 | 9.09 | 5.94 | 8.94 | -1.74 | 27.33 | 0.3 | -0.69 | 0.24 |
| | state 2 | 141 | 15.34 | 7.11 | 15.29 | -1.04 | 36.52 | 0.54 | 0.27 | 0.6 |
| | All states | 756 | 10.25 | 6.64 | 9.99 | -1.74 | 36.52 | 0.52 | 0.2 | 0.24 |
| Workers-Compen | state 1 | | | | | | | | | |


Rho
-1.0 -0.8 -0.6 -0.4 -0.2 0.0 0.2 0.4 0.6 0.8 1.0
LGC map - State 1

## Densities, Scatter plot and correlation of true states



# Test of asymmetric dependence between states

Hypothesis:

$$H_0 : \quad \rho_1(x_i, y_j) =$$

# Conclusion

- Complex dependency structures can be modeled using a mix of LGC and HMMs

# Conclusion

# Conclusion

- Complex dependency structures can be modeled using a mix of LGC and HMMs
- Model found time carrying dependency-financial crisis
- Dependency structure differs

# Conclusion

- Complex dependency structures can be modeled using a mix of LGC and HMMs
- Model found time carrying dependency-financial crisis
- Dependency structure differs
- This analysis enables insurers to better assess and manage their overall risk exposure across different lines of business especially during economic,political crisis periods

# Further work

- Try t-distribution instead of normal for the HMMs

# Further work

- Try t-distribution instead of normal for the HMMs
- consider the other data set (content + house insurance from Norway) + publicly available data

# Further work

- Try t-distribution instead of normal for the HMMs
- consider the other data set (content + house insurance from Norway) + publicly available data
- Compare with other dependency modeling techniques like Bernstein copulas, Vine Copulas?

# Further work

- Try t-distribution instead of normal for the HMMs
- consider the other data set (content + house insurance from Norway) + publicly available data
- Compare with other dependency modeling techniques like Bernstein copulas, Vine Copulas?
- Include covariates? claim reporting delays, settlement delays,

# Further work

- Try t-distribution instead of normal for the HMMs
- consider the other data set (content + house insurance from Norway) + publicly available data
- Compare with other dependency modeling techniques like Bernstein copulas, Vine Copulas?
- Include covariates? claim reporting delays, settlement delays,
- Apply to Reserving, pricing,reinsurance arrangements?