Insurance Claims Modulated by a Hidden Marked Point Process

Robert J. Elliott *
Haskayne School of Business
University of Calgary
Calgary, Alberta, Canada

*This is a joint with Tak Kuen Siu and Hailiang Yang. Robert Elliott would like to thank SSHRC for its continued support. Hailiang Yang would like to acknowledge the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. 7426/06H).
Outline of the presentation:

- Background
- Model dynamics and change of measures
- Filters and smoothers
- Markov-switching stochastic intensity and claim sizes
- Parameter estimation: EM algorithm
- Summary
§1. Background

1.1. Compound Poisson process for actuarial use

- Compound Poisson process: Standard and classical model for insurance claims in ruin theory.

- Describe aggregate insurance claims over a fixed time horizon:
  1. The number of claims modeled as a Poisson process.
  2. The amounts of individual claims modeled as a sequence of positive random variables.
3. Claim frequency and claim amounts are independent.

- Popular: Analytically tractable results for ruin probabilities.

- References: Rolski et al. (1999) and Assmussan (2000).
1.2. Regime-switching models for actuarial use


- Include a continuous-time Markov chain whose states represent different economic environments.

- Model the change in the state of an economy due to structural changes in the (macro)-economic conditions and business cycles.

- Observable Claim frequency and claim sizes.
• Unobservable Markov chain $\implies$ Model uncertainty.

• Two key problems in the implementation of the Markov-modulated compound Poisson model:

  1. How to estimate the hidden risk state?

  2. How to estimate the parameters of the model?
1.3. Key points of our work

- Develop methods for filtering and smoothing the hidden states of Markov-modulated compound Poisson processes based on the observed information for the number of claims and the claim sizes.

- Case I: Stochastic intensity switches over time according to a continuous-time hidden Markov chain.

- Case II: Both the stochastic intensity and the distribution of the claim sizes depend on the hidden Markov chain.
• Derive robust filters and smoothers in the form of O.D.E.s in both cases.

• Estimate the model parameters using the robust filter-based and smoother-based EM algorithms.
§2. Model Dynamics and Change of Measures

• Consider a Markov-modulated compound Poisson model with a Markov-switching stochastic intensity only for aggregate insurance claims.

• Describe the hidden states of an economy by the states of a continuous-time hidden Markov chain.

• \( \{X_t\}_{t \in T} \): A continuous-time hidden Markov chain on \((\Omega, \mathcal{F}, \mathcal{P})\) with state space \(\{e_1, e_2, \ldots, e_N\}\), a finite set of unit vectors with \(e_i = (0, \ldots, 1, \ldots, 0) \in \mathbb{R}^N\).
• **A**: The generator or the rate matrix \([a_{ij}]_{i,j=1,2,...,N}\).

• Semi-martingale decomposition by Elliott et al. (1994):

\[
X_t = X_0 + \int_0^t AX_s ds + M_t,
\]

where \(\{M_t\}_{t \in T}\) is a \(\mathbb{R}^N\)-valued martingale with respect to \((\mathcal{F}^X, \mathcal{P})\).
Consider a Poisson process $N := \{N_t\}_{t \in \mathcal{T}}$ on $(\Omega, \mathcal{F}, \mathcal{P})$, whose stochastic intensity is:

$$\lambda_t := \langle \lambda, X_t \rangle = \sum_{i=1}^{K} \langle \lambda, e_i \rangle I_{\{X_t = e_i\}},$$

where $\lambda := (\lambda_1, \lambda_2, \ldots, \lambda_K) \in \mathbb{R}^K$ and $\lambda_k \geq 0$, for each $k = 1, 2, \ldots, K$.

- $N_t$: the number of claims over the time $[0, t]$.

- Consider right-continuous, complete versions of the filtrations

$$\mathcal{F}_t^X := \{\mathcal{F}_t^X\}_{t \in \mathcal{T}}, \quad \mathcal{F}_t^X := \sigma\{X_u | u \in [0, t]\}, \quad \mathcal{F}_t^N := \sigma\{N_u | u \in [0, t]\}, \quad \mathcal{G}_t := \{\mathcal{G}_t\}_{t \in \mathcal{T}}, \quad \mathcal{G}_t := \mathcal{F}_t^X \vee \mathcal{F}_t^N.$$
The Doob-Meyer decomposition for $N$ is (see Elliott and Malcolm (2005)):

$$N_t = \int_0^t \langle \lambda, X_u \rangle \, du + V_t,$$

where $V := \{V_t\}_{t \in T}$ is a $(\mathcal{P}, \sigma\{G_u | u \in [0, t]\})$-martingale.

Define a probability distribution $F_Y(\cdot)$ on $(\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+))$, where $F_Y(\cdot)$ is given.

Suppose the total amount of claims to time $t$ is:

$$Z_t = \int_0^t \int_0^\infty y dF_Y(y) dN_u$$

The jump sizes provide no extra information about $X$. 
• Consider a reference probability measure \( \mathcal{P}^\dagger \) under which \( N \) is a Poisson process with unit intensity and is independent of \( X \).

• Under \( \mathcal{P}^\dagger \),

\[
Q_t := N_t - t ,
\]

is a martingale.
• Define a process \( \Lambda := \{\Lambda_{0,t}\}_{t \in T} \):

\[
\Lambda_{0,t} := \prod_{0 < u \leq t} \langle X_u, \lambda \rangle^\Delta N_u \exp \left( \int_0^t (1 - \langle X_u, \lambda \rangle) du \right)
\]

\[
= 1 + \int_0^t \Lambda_{0,u-} (\langle X_u-, \lambda \rangle - 1) dQ_u
\]

• \( \Lambda \) is a \((G, P^\dagger)\)-martingale.

• Define the real-world probability measure \( P \) by setting

\[
\Lambda_{0,t} := \frac{dP}{dP^\dagger} \bigg|_{G_t}
\]

• Suppose \( \gamma := \{\gamma_t\}_{t \in T} \) is any \( G \)-adapted process.
• Given $\mathcal{F}_t^N$, estimate $\gamma_t$ by its least-square estimate $E[\gamma_t|\mathcal{F}_t^N]$.

• By a form of Bayes’ rule (see Elliott et al. (1994)),

$$E[\gamma_t|\mathcal{F}_t^N] = \frac{E^\dagger[\Lambda_{0,t}\gamma_t|\mathcal{F}_t^N]}{E^\dagger[\Lambda_{0,t}|\mathcal{F}_t^N]} = \frac{\sigma_t(\gamma)}{\sigma_t(1)},$$

where $E^\dagger$ denotes expectation with respect to $\mathcal{P}^\dagger$. 
§3. Filters and Smoothers

- Derive the filters and the smoothers for $X$.

- Filters: Derive an O.D.E. satisfied by the transformed process of the filtered estimates.

- Smoothers: Derive linear forward and backward O.D.E.s for the smoothed estimates and the process of extra information, respectively.

- Advantage: Compute the filters and smoothers without recourse to stochastic integration.
3.1. Filtered estimates

- Define $q_t := E^t[\Lambda_{0,t} X_t | \mathcal{F}_t^N] \in \mathbb{R}^K$, the unnormalized conditional distribution of $X$ given the observations.

- **Theorem 3.1:** $q_t$ satisfies the S.D.E.:
  
  $$q_t = q_0 + \int_0^t A q_u du + \int_0^t \text{diag}\{\langle \lambda, e_k \rangle - 1\} q_u - dQ_u$$

- By the Bayes rule and noting that $\langle X_t, 1 \rangle = 1$,
  
  $$p_t := E[X_t | \mathcal{F}_t^N] = \frac{q_t}{\langle q_t, 1 \rangle},$$

  where $1 := (1, 1, \ldots, 1) \in \mathbb{R}^K$. 

• Define a matrix-valued stochastic process \( \{ \Gamma_t \} \) such that \( \Gamma_t \in \mathbb{R}^{K \times K}, t \in T \), and

\[
\Gamma_t := \text{diag}\{\gamma_1^t, \gamma_2^t, \ldots, \gamma_K^t\},
\]

where \( \gamma_k^t := \exp[(1 - \langle \lambda, e_k \rangle)t] \langle \lambda, e_k \rangle^{N_t}, k = 1, 2, \ldots, K \).

• Define a transformed process \( \bar{q}_t := \Gamma_t^{-1} q_t \).

• \textbf{Theorem 3.2:} \( \bar{q}_t \) satisfies the forward linear O.D.E.:

\[
\frac{d\bar{q}_t}{dt} = \Gamma_t^{-1} A \Gamma_t \bar{q}_t, \quad \bar{q}_0 = q_0 \in \mathbb{R}^K
\]
Lemma 3.3: Let

\[ \pi(X_t) := \frac{\Gamma_t \bar{q}_t}{\langle \Gamma_t \bar{q}_t, 1 \rangle} . \]

Then, \( \pi(X_t) \) is a version of \( E[X_t|\mathcal{F}_t^N] \), which is continuous in the observation process \( N \) in the Skorokhod topology.
3.2. Smoothed estimates

- Evaluate the conditional expectation $E[X_t|\mathcal{F}_T^N]$, for $t \in [0, T]$.

- By the Bayes rule,
  \[
  E[X_t|\mathcal{F}_T^N] = \frac{E^\dagger[\Lambda_{0,T}X_t|\mathcal{F}_T^N]}{E^\dagger[\Lambda_{0,T}|\mathcal{F}_T^N]}
  \]

- Let $r_t := E^\dagger[\Lambda_{0,T}X_t|\mathcal{F}_T^N]$, which is the unnormalized smoother of $X_t$ given $\mathcal{F}_T^N$.

- By the semi-group property of $\Lambda$ and the double expectation:
  \[
  r_t = E^\dagger[\Lambda_{0,t}X_tE^\dagger[\Lambda_{t,T}|\mathcal{F}_T^N \vee \mathcal{F}_t^X]|\mathcal{F}_T^N]
  \]
• Due to the Markov property of $X$ under $\mathcal{P}^\dagger$,

$$E^\dagger[\Lambda_{t,T}|\mathcal{F}_T^N \vee \mathcal{F}_t^X] = E^\dagger[\Lambda_{t,T}|\mathcal{F}_T^N \vee \sigma\{X_t\}]$$

• Define a process $\nu$:

$$\nu^k_t := E^\dagger[\Lambda_{t,T}|\mathcal{F}_T^N, X_t = e_k] ,$$

and $\nu_t := (\nu^1_t, \nu^2_t, \ldots, \nu^K_t) \in \mathbb{R}^K$.

• Note that $\sum_{k=1}^K \langle X_t, e_k \rangle = 1$, that $\nu^k_t$ is $\mathcal{F}_T^N$-measurable, and that $N$ has independent increments under $\mathcal{P}^\dagger$.

• Then,

$$r_t = \sum_{k=1}^K q^k_t \nu^k_t e_k \in \mathbb{R}^K$$
• The normalized smoothed estimate of $X$:

$$p_t := E[X_t|\mathcal{F}_T^N] = \frac{r_t}{\langle r_t, 1 \rangle}$$

• The denominator of the normalized smoothed estimate:

$$\langle r_t, 1 \rangle = \langle q_t, r_t \rangle$$

• Note that the process $\langle r_t, 1 \rangle$ is independent of time, so is $\langle q_t, \nu_t \rangle$.

• Then,

$$\frac{d}{dt} \langle r_t, 1 \rangle = \frac{d}{dt} \langle q_t, \nu_t \rangle = 0$$
• **Theorem 3.4:** For $t \in [0, T],

\[
p_t = \frac{1}{\langle \bar{q}_t, \bar{\nu}_t \rangle} \sum_{k=1}^{K} \langle \bar{q}_t, e_k \rangle \langle \bar{\nu}_t, e_k \rangle e_k,
\]

where $\bar{q}_t$ satisfies:

\[
\frac{d\bar{q}_t}{dt} = \Gamma^{-1}_t A \Gamma_t \bar{q}_t, \quad \bar{q}_0 = q_0,
\]

and $\bar{\nu}_t$ satisfies:

\[
\frac{d\bar{\nu}_t}{dt} = -\Gamma_t A^* \Gamma^{-1}_t \bar{\nu}_t, \quad \bar{\nu}_T = \Gamma_T 1
\]
§4. Markov-Switching Stochastic Intensity and Claim Sizes

- Consider a Markov-modulated marked point process \( Z := \{Z_t\}_{t \in \mathcal{T}} \) under \( \mathcal{P} \).

- \( \mathcal{X} \): the product space \( \mathcal{T} \times \mathcal{Z} \), where \( \mathcal{T} := [0, T] \) and \( \mathcal{Z} := (0, \infty) \).

- \( \gamma(\cdot, \cdot) \): A random measure on \( \mathcal{X} \).
\* \* \* \* \* \* \*

- γ is a sum of random delta functions:

\[ \gamma(dy, dt; \omega) = \sum_{k} \delta(Y_{T_k}(\omega))\delta(T_k(\omega)) \]

so that for suitable integrands \( f : (\Omega \times (0, \infty) \times [0, \infty)) \rightarrow \mathbb{R} \),

\[ \int_{0}^{t} \int_{0}^{\infty} f(\omega, y, u)\gamma(dy, du) = \sum_{T_k \leq t} f(\omega, Y_{T_k}(\omega), T_k(\omega)) \]

- Assume \( Z \) follows:

\[ Z_t = \int_{0}^{t} \int_{0}^{\infty} y\gamma(dy, du) \]

- \( f_k(y) \): the probability density function of the random claim size \( y := Z_u - Z_{u-} \) when \( X_{u-} = e_k, k = 1, 2, \ldots, K \).
• The number of claim arrivals, $N_t$, over $[0,t]$ is a Poisson random variable with stochastic intensity:

$$\lambda_t := \langle \lambda, X_t \rangle,$$

where $\lambda := (\lambda_1, \lambda_2, \ldots, \lambda_K) \in \mathbb{R}^K$.

• Conditional on $X$, the times of claim arrivals and the claim sizes are independent.

• The compensator of $\gamma(dy, du)$ conditional on $X_{u-}$ under $\mathcal{P}$:

$$\nu(dy, du|X_{u-}) := \sum_{k=1}^{K} \langle X_{u-}, e_k \rangle \lambda_k f_k(y) dy du$$

• $\bar{G}_t := \mathcal{F}_t^Z \vee \mathcal{F}_t^X$. 
• Under $\mathcal{P}$,

$$\tilde{M}_t := Z_t - \int_0^t \int_0^\infty y\nu(dy, du|X_{u-}) ,$$

is a $(\bar{G}, \mathcal{P})$-local martingale.

• Assume that under a reference probability measure $\mathcal{P}^\dagger$, $Z$ is a compound Poisson process with unit intensity and a density function for the claim sizes $f(y)$.

• Under $\mathcal{P}^\dagger$, the compensator $\nu^\dagger$ of $\gamma$ is:

$$\nu^\dagger(dy, du) := f(y)dydu$$
• Under $\mathcal{P}^\dagger$,

$$M^\dagger_t := Z_t - \int_0^t \int_0^\infty y\nu^\dagger(dy, du),$$

is a local martingale under $\mathcal{P}^\dagger$.

• $h_k(y) = \frac{\lambda_k f_k(y)}{f(y)}$, $k = 1, 2, \ldots, K$.

• Define a density process $\bar{\Lambda} := \{\bar{\Lambda}_{0,t}\}_{t \in \mathcal{T}}$ giving the change of measure:

$$\bar{\Lambda}_{0,t} = \exp \left( -\int_0^t \sum_{k=1}^K \langle X_{u-}, e_k \rangle \int_0^\infty (h_k(y) - 1) f(y) dy du 
+ \int_0^t \sum_{k=1}^K \langle X_{u-}, e_k \rangle \int_0^\infty \log h_k(y) \gamma(dy, du) \right)$$
• $\tilde{\Lambda}$ is a $(\bar{G}, \mathcal{P}^\dagger)$-local martingale and assume it is a $(\bar{G}, \mathcal{P}^\dagger)$-martingale.

• Define the real-world probability measure $\mathcal{P}$ by setting:

$$\Lambda_{0,t} := \left. \frac{d\mathcal{P}}{d\mathcal{P}^\dagger} \right|_{\bar{G}_t}$$

• By Girsanov’s theorem for jump processes in Elliott (1982), $\hat{M}$ is a $(\bar{G}, \mathcal{P})$-local martingale.
• Derive a recursive Zakai equation for the filter $E(X_t | \mathcal{F}_t^Z)$, where $\mathcal{F}_t^Z = \sigma\{Z_u | u \in [0, t]\}$.

• Suppose $Y_u : \Omega \rightarrow (0, \infty)$ is a random variable with $Y_u(\omega) > 0$ and density function $f$ under $\mathcal{P}^\dagger$.

• Define

$$H_k(u, \omega) := \frac{\lambda_k f_k(Y_u(\omega))}{f(Y_u(\omega))} = h_k(Y_u(\omega))$$

• Consider the diagonal matrices:

$$\text{diag}(H - 1) := \text{diag}(H_1(u, \omega) - 1, \ldots, H_K(u, \omega) - 1),$$
$$\text{diag}(\lambda - 1) := \text{diag}(\lambda_1 - 1, \ldots, \lambda_K - 1)$$
• **Theorem 4.1:** Let $q_t = \sigma(X_t) = E^t(\bar{\Lambda}_{0,t}X_t|\mathcal{F}_t^Z)$. Then,

$$q_t = q_0 + \int_0^t Aq_u du + \int_0^t \text{diag}(H - 1)q_u dN_u$$

$$- \int_0^t \text{diag}(\lambda - 1)q_u du$$

• Derive a robust filter for $X$.

• Consider the process:

$$\gamma_t^k := \exp\left[(1 - \lambda_k)t + \int_0^t \log H_k(u, \omega) dN_u\right]$$

• Define $\Gamma_t := \text{diag}(\gamma_t^1, \gamma_t^2, \ldots, \gamma_t^K)$. 
• **Theorem 4.2**: Let $\bar{q}_t := \Gamma_t^{-1} q_t$. Then, $\bar{q}$ satisfies the linear O.D.E.:

$$\bar{q}_t = q_0 + \int_0^t \Gamma_u^{-1} A \Gamma_u \bar{q}_u du$$

• Note that $Z$ has independent increments under $\mathcal{P}^\dagger$.

• Repeat the same procedure as in 3.2 to derive $p_t := E(X_t|\mathcal{F}_T^Z)$.

• $p_t$ is given by Theorem 3.4 with $\mathcal{F}_t^N$ replaced by $\mathcal{F}_t^Z$. 
§5. Parameter Estimation by the EM Algorithm

- Estimate the parameters in the Markov-modulated marked point process with Markov-switching stochastic intensity and claim sizes using the EM algorithm.


- Compute the estimates for \( A := [a_{ij}]_{i,j=1,2,...,K} \) and \( \lambda := (\lambda_1, \lambda_2, ..., \lambda_K) \).
• Assume that the distribution of claim sizes $F_k(y)$ is known, for each regime $k = 1, 2, \ldots, K$.

• Estimate the distributions $F_1(y), F_2(y), \ldots F_K(y)$:

  1. Divide a given set of claims data into $K$ groups using some simple criteria.

  2. For example, determine a set of threshold parameters, say $R_1 < R_2 < \cdots < R_{K-1}$, and allocate the claim data $Y_t$ into the $k^{th}$ group if $Y_t \in [R_{k-1}, R_k]$.

  3. Estimate $F_k(y)$ using the claims data in the $k^{th}$ group using the method outlined in Klugman et al. (2004).
• Develop a robust filter-based EM algorithm and a robust smoother-based EM algorithm.

• Provide practical forms of the robust dynamics in the estimation scheme by computing time domain discretization of these robust dynamics.
5.1. Robust filter-based EM algorithm

- Since $F_k(y)$ is supposed to be given, the observation processes $N$ and $Z$ provide the same amount of information to estimate $a_{ij}$ and $\lambda_i$.

- The estimators $\hat{a}_{ij}$ and $\hat{\lambda}_i$ are (see Dembo and Zeitouni (1986)):

$$
\hat{a}_{ij} = \frac{E[N_{ij}^{i}|\mathcal{F}_T^N]}{E[O_{ij}^{i}|\mathcal{F}_T^N]} = \frac{E[N_{ij}^{i}|\mathcal{F}_T^Z]}{E[O_{ij}^{i}|\mathcal{F}_T^Z]},
$$

and

$$
\hat{\lambda}_i = \frac{E[G_i^i|\mathcal{F}_T^N]}{E[O_i^i|\mathcal{F}_T^N]} = \frac{E[G_i^i|\mathcal{F}_T^Z]}{E[O_i^i|\mathcal{F}_T^Z]}.
$$
• For any $\bar{\gamma}$-adapted integrable process $\gamma := \{\gamma_t\}_{t \in T}$,

$$\sigma(\gamma_t) := E^\dagger[\bar{\Lambda}_0, t \gamma_t | \mathcal{F}_t^Z]$$

• Then,

$$\hat{a}_{ij} = \frac{\sigma(N_{ij}^T)}{\sigma(O_{ij}^T)} , \quad \hat{\lambda}_i = \frac{\sigma(G_{ii}^T)}{\sigma(O_{ii}^T)}$$

• Define the following quantities:

$$O_t^i := \int_0^t \langle X_u, e_i \rangle \, du \in \mathbb{R} ,$$

$$N_{ij}^T := \int_0^t \langle X_u-, e_i \rangle \langle dX_u, e_j \rangle \, du \in \mathbb{R} ,$$
and

\[ G_t^i := \int_0^t \langle X_u, e_i \rangle \, dN_u \in \mathbb{R} \]

- Write \( \lambda - 1 := (\lambda_1 - 1, \ldots, \lambda_K - 1) \) and \( h := (h_1(y), \ldots, h_K(y)) \).
The dynamics for the measure-valued quantities $\sigma(N_{ij}^t X_t)$, $\sigma(O_i^t X_t)$ and $\sigma(G_i^t X_t)$ are:

\[
\sigma(G_i^t X_t) = \int_0^t A\sigma(G_u^i X_u) du + \int_0^t \text{diag}(H - 1)\sigma(G_u^i X_{u-}) dN_u \\
- \int_0^t \text{diag}(\lambda - 1)\sigma(G_u^i X_u) du \\
+ \int_0^t \langle h, e_i \rangle \langle q_{u-}, e_i \rangle dN_u e_i \\
- \int_0^t \langle \lambda - 1, e_i \rangle \langle q_u, e_i \rangle du e_i ,
\]
\[
\sigma(N^i_{t} X_t) = \int_{0}^{t} A\sigma(N^{ij}_{u} X_u) du + \int_{0}^{t} \langle q_{u-}, e_i \rangle \langle Ae_i, e_j \rangle du e_j \\
+ \int_{0}^{t} \text{diag}(H - 1)\sigma(N^{ij}_{u-} X_{u-}) dN_u \\
- \int_{0}^{t} \text{diag}(\lambda - 1)\sigma(N^{ij}_{u} X_u) du ,
\]

\[
\sigma(O^{i}_{t} X_t) = \int_{0}^{t} A\sigma(O^{i}_{u} X_u) du + \int_{0}^{t} \langle q_{u}, e_i \rangle du e_i \\
+ \int_{0}^{t} \text{diag}(H - 1)\sigma(O^{i}_{u-} X_{u-}) dN_u \\
- \int_{0}^{t} \text{diag}(\lambda - 1)\sigma(O^{i}_{u} X_u) du
\]
• Note that

\[
\langle \sigma(N_{t}^{ij}X_{t}), 1 \rangle = \langle E^{\dagger}[\Lambda_{0,t}N_{t}^{ij}X_{t}], 1 \rangle = \sigma(N_{t}^{ij}) , \\
\langle \sigma(O_{t}^{i}X_{t}), 1 \rangle = \sigma(O_{t}^{i}) , \\
\langle \sigma(G_{t}^{i}X_{t}), 1 \rangle = \sigma(G_{t}^{i})
\]

• Hence,

\[
\hat{a}_{ij} = \frac{\langle \sigma(N_{T}^{ij}X_{T}), 1 \rangle}{\langle \sigma(O_{T}^{ij}X_{T}), 1 \rangle} , \\
\hat{\lambda}_{i} = \frac{\langle \sigma(G_{T}^{i}X_{T}), 1 \rangle}{\langle \sigma(O_{T}^{i}X_{T}), 1 \rangle}
\]

• Implement a filter bank consisting of recursive filters involving stochastic integrals to compute the estimators.
• Use a version of a gauge transformation introduced by Clark (1978) to eliminate stochastic integrals and to develop robust filters which do not involve stochastic integrals.

• Derive robust filtering equations corresponding to \( \sigma(G_t^i X_t) \), \( \sigma(N_t^{ij} X_t) \) and \( \sigma(O_t^i X_t) \).

• Define \( \Gamma_t := \text{diag}(\gamma_1^t, \gamma_2^t, \ldots, \gamma_K^t) \), where \( \gamma_k^t := \exp(X_k^t) \) and

\[
X_k^t := (1 - \lambda_k) t + \int_0^t \log H_k(u, \omega) dN_u
\]

• Write \( \bar{\sigma}(G_t^i X_t) := \Gamma_t^{-1} \sigma(G_t^i X_t) \).
\[ \bar{\sigma}(G_t^i X_t) = \int_0^t \Gamma_u^{-1} A \Gamma_u \bar{\sigma}(G_u^i X_u) du + \text{diag} \left( \frac{1}{H} \right) \langle h, e_i \rangle N_t e_i \]

\[ - \int_0^t N_u \langle h, e_i \rangle \left\langle d \left[ \text{diag} \left( \frac{1}{H} \right) \bar{q}_u \right], e_i \right\rangle e_i , \]

\[ \bar{\sigma}(N_t^{ij} X_t) = \int_0^t \Gamma_u^{-1} A \Gamma_u \bar{\sigma}(N_u^{ij} X_u) du + \int_0^t \langle \bar{q}_u, e_i \rangle \langle A e_i, e_j \rangle du e_j , \]

\[ \bar{\sigma}(O_t^i X_t) = \int_0^t \Gamma_u^{-1} A \Gamma_u \bar{\sigma}(O_u^i X_u) du + \int_0^t \langle \bar{q}_u, e_i \rangle du e_i , \]
\[ \ddot{q}_t = q_0 + \int_0^t \Gamma_u^{-1} A \Gamma_u \ddot{q}_u du , \]

where

\[ \text{diag}\left( \frac{1}{H} \right) = \text{diag}\left( \frac{1}{H_k(t, \omega)}, \ldots, \frac{1}{H_K(t, \omega)} \right) \]

- **Practical Implementation**: Time discretization of the robust filtering equations.

- **Time discretization of \( q_t \):**

\[ q_{tm} \approx \Phi_m[I + A\Delta]q_{tm-1} , \]

where \( \Phi_m := \Gamma_{tm} \Gamma_{tm}^{-1} \).
• Time discretization of $\sigma(G^i_t X_t)$:

$$
\sigma(G^i_{tm} X_{tm}) 
\approx \sigma(G^i_{tm-1} X_{tm-1}) + \Gamma_{tm-1} \langle h, e_i \rangle \left[ \text{diag} \left( \frac{1}{H_{tm}} \right) \langle \bar{q}_{tm}, e_i \rangle N_{tm} e_i 
- \text{diag} \left( \frac{1}{H_{tm-1}} \langle \bar{q}_{tm-1}, e_i \rangle N_{tm-1} e_i \right) \right] 
- \Gamma_{tm-1} N_{tm} \langle h, e_i \rangle \left[ \text{diag} \left( \frac{1}{H_{tm}} \right) \bar{q}_{tm}, e_i \right] e_i 
- \Gamma_{tm-1} N_{tm-1} \langle h, e_i \rangle 
\left[ \text{diag} \left( \frac{1}{H_{tm-1}} \right) \Gamma_{tm-1}^{-1} (A - I) q_{tm-1}, e_i \right] \Delta e_i
$$
• Time discretization of $\sigma(N_{t}^{ij} X_t)$:

$$\sigma(N_{t_m}^{ij} X_{t_m}) \approx \Phi_m[I + \Delta A] \sigma(N_{t_{m-1}}^{ij}) + \Phi_m \langle q_{t_{m-1}}, e_i \rangle$$

$$+ \Delta \langle Ae_i, e_j \rangle \Delta e_i$$

• Time discretization of $\sigma(O_{t}^{i} X_t)$:

$$\sigma(O_{t_m}^{i} X_{t_m}) \approx \Phi_m[I + \Delta A] \sigma(O_{t_{m-1}}^{i} X_{t_{m-1}}) + \Phi_m \langle q_{t_{m-1}}, e_i \rangle e_i$$
• The estimators \( \hat{a}_{ij} \) and \( \hat{\lambda}_i \) can be computed by the three steps of the filter-based EM algorithm.

1. **Step I:** Select the initial values \( \hat{a}_{ij}(0) \) and \( \lambda_i(0) \).

2. **Step II:** Compute the MLEs

\[
\hat{a}_{ij} = \frac{\langle \sigma(N^i_T X_T), 1 \rangle}{\langle \sigma(O^i_T X_T), 1 \rangle}, \quad \hat{\lambda}_i = \frac{\langle \sigma(G^i_T X_T), 1 \rangle}{\langle \sigma(O^i_T X_T), 1 \rangle}
\]

3. **Step III:** Stop or continue from Step II.
5.2. Smoother-based EM algorithm

- Compute estimators based on smoothing schemes rather than filtering schemes.

- Useful when the expectation step is completed with smoothed estimates, rather than filtered estimates, in some practical implementations of the EM algorithm.

- Main difficulty: Difficult to develop the backwards dynamics, which involve the construction of stochastic integrals evolving backwards in time.
• Use a duality between forward and backwards robust dynamics to develop smoothing algorithms, which do not involve stochastic integrals at all.

• Compute dynamics for the dual process \( \tilde{\nu} \) for \( \nu \), where \( \nu_t := (\nu^1_t, \nu^2_t, \ldots, \nu^K_t) \in \mathbb{R}^K \) with

\[
\nu^k_t := E^\dagger [\Lambda_{t,T}|\mathcal{F}_T^N, X_t = e_k]
\]

• Find a process \( \tilde{\nu} \) such that the duality holds:

\[
\langle \bar{q}_t, \tilde{\nu}_t \rangle = \langle \Gamma_t^{-1}q_t, \Gamma_t\nu_t \rangle = \langle q_t, \nu_t \rangle, \forall t \in T
\]
• Let $\bar{\nu}_t := \Gamma_t \nu_t$. By Theorem 3.4,

$$\frac{d\bar{\nu}_t}{dt} = -\Gamma_t A^* \Gamma_t^{-1} \bar{\nu}_t,$$

where $\bar{\nu}_T = \Gamma_T \nu_T = \Gamma_T 1$.

• By exploiting the duality,

$$\left\langle \sigma(G_T^i X_T), \nu_T \right\rangle$$

$$= \left\langle \bar{\sigma}(G_T^i X_T), \bar{\nu}_T \right\rangle$$

$$= \text{diag} \left( \frac{1}{H} \right) N_T \left\langle h, e_i \right\rangle \left\langle \bar{q}_T, e_i \right\rangle \left\langle e_i, \bar{q}_T \right\rangle$$

$$- \int_0^T N_t \left\langle h, e_i \right\rangle \left\langle e_i, \bar{\nu}_t \right\rangle \left\langle d \left[ \text{diag} \left( \frac{1}{H} \right) \bar{q}_t \right], e_i \right\rangle$$
Similarly,

\[
\langle \sigma(N_{ij}^T X_T), \nu_T \rangle = \langle \bar{\sigma}(N_{ij}^T X_T), \bar{\nu}_T \rangle \\
= \int_0^T \langle Ae_i, e_j \rangle \langle q_t, e_i \rangle \langle \nu_t, e_j \rangle \, dt,
\]

\[
\langle \sigma(O_i^T X_T), \nu_T \rangle = \langle \bar{\sigma}(O_i^T X_T), \bar{\nu}_T \rangle \\
= \int_0^T \langle q_t, e_i \rangle \langle \nu_t, e_i \rangle \, dt
\]
The smoother-based update equations are:

\[
\hat{a}_{ij}(k + 1) = \hat{a}_{ij}(k) \frac{\int_0^T \langle q_t, e_i \rangle \langle \nu_t, e_j \rangle \, dt}{\int_0^T \langle q_t, e_i \rangle \langle \nu_t, e_i \rangle \, dt},
\]

and

\[
\hat{\lambda}_i(k + 1) = \left[\text{diag}\left(\frac{1}{H}\right)N_T \langle h, e_i \rangle \langle \overline{q}_T, e_i \rangle \langle e_i, \overline{q}_T \rangle - \int_0^T N_t \langle h, e_i \rangle \langle e_i, \overline{\nu}_t \rangle \left\langle d\left[\text{diag}\left(\frac{1}{H}\right)\overline{q}_t\right], e_i \right\rangle \right]
\left(\int_0^T \langle q_t, e_i \rangle \langle \nu_t, e_i \rangle \, dt\right)^{-1}
\]
These estimates can be computed by following the three steps in the smoother-based EM algorithm:

1. **Step I:** Select $\hat{a}_{ij}(0)$ and $\hat{\lambda}_i(0)$.

2. **Step II:** Compute the MLEs, $\hat{a}_{ij}(k + 1)$ and $\hat{\lambda}_i(k + 1)$, respectively.

3. **Step III:** Stop or continue from Step II.
§6. Summary

- Developed a way to filter and smooth Markov-modulated compound Poisson model and marked point process for actuarial use and other potential applications.

- Considered the case that the stochastic intensity switches over time according to the state of a continuous-time finite-state hidden Markov chain.

- Considered the general case that both the stochastic intensity and the distribution of the jump size depend on the state of the hidden Markov chain.
• Derived robust filters and smoothers in the form of O.D.E.s in both cases, which provide a method to select or estimate risk model in the “mean-square-error” sense.

• Provided methods to estimate model parameters of a Markov-modulated marked point process based on the robust filter-based and smoother-based EM algorithms

~ Thank you! ~
References


