Two-State Imprecise Markov Chains for Statistical Modelling of Two-State Non-Markovian Processes

Matthias C. M. Troffaes & Thomas Krak & Henna Bains

Durham University, United Kingdom & Gent University, Belgium

4 July, 2019

Aim

present a new framework for

- fitting an imprecise two-state Markov chain
- to data generated by a two-state non-Markovian process
- via an imprecise version of MCMC

The Old Way

1. transition times $T_i \sim \mathrm{Exp}(\lambda_i)$
   (i.e. precise Markov chain assumed)
2. set of priors for λ_i
3. combine with data to get set of posteriors for λ_i
4. posterior predictive bounds on λ_i
5. use these bounds to fix an imprecise Markov chain

what is wrong with this approach?

The Old Way

what is wrong with this approach?

- model = set of distributions on parameters of a precise Markov chain
  imprecise Markov chains are not equivalent to this model
- sampling uncertainty ignored: only posterior predictive bounds are used
  no full uncertainty quantification
- imprecision does not reflect violations of stationarity & Markovianity
  wasted opportunity
- non-Bayesian analysis: inferences are not coherent with the model

The Newly Proposed Way

1. transition times $T_i \sim f_i(\lambda_i)$
   where f_i is a non-Markovian and/or non-stationary process
   and λ_i are unknown parameters of this process
2. set of prior distributions on λ_i
3. for each prior, use standard MCMC to sample posterior realizations of λ_i
   i.e. fit the non-Markovian process to the data
4. for each posterior sample of λ_i bound process by an imprecise Markov chain
5. produce sample of posterior bounds for any predictive quantity

The Newly Proposed Way

Nice features of this new approach:

- a form of imprecise MCMC: imprecision built into the predictive part
  this works because of Walley's marginal extension theorem
- fully coherent robust Bayesian approach:
  predictions are directly derived from posterior distribution
- imprecision reflects lack of data & non-Markovianity & non-stationarity
  (even if there is lot's of data, inferences can remain imprecise)

How Did We Do It

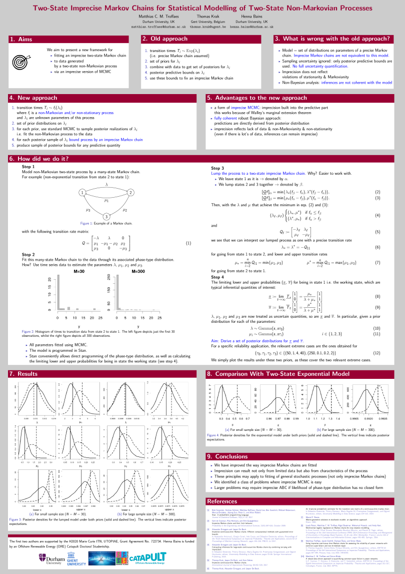
- non-Markovian behaviour modelled by a multi-state Markov chain
  can model any phase-type distribution for transition times!
  (every distribution can be approximated by a phase-type distribution)
- bounding by lumping multi-state Markov chains into imprecise two-state Markov chains
- details: see poster

Conclusions

- we've improved the way imprecise Markov chains are fitted
- imprecision results not only from limited data but also from characteristics of the process
- principles may apply to fitting of general stochastic processes
  (not only imprecise Markov chains)
- we identified a class of problems where imprecise MCMC is easy
- larger problems may require imprecise ABC
  if likelihood of phase-type distribution has no closed form

Thank You For Your Attention

We look forward to see you at our poster!



References I

[1] Bob Carpenter, Andrew Gelman, Matthew Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell.
Stan: A probabilistic programming language.
Journal of Statistical Software, 76(1):1–32, 2017.

[2] Gert de Cooman, Filip Hermans, and Erik Quaeghebeur.
Imprecise Markov chains and their limit behavior.
Probability in the Engineering and Informational Sciences, 23(4):597–635, October 2009.

[3] Alexander Erreygers and Jasper De Bock.
Imprecise continuous-time Markov chains: Efficient computational methods with guaranteed error bounds.
In Alessandro Antonucci, Giorgio Corani, Inés Couso, and Sébastien Destercke, editors, *Proceedings of the Tenth International Symposium on Imprecise Probability: Theories and Applications*, volume 62 of *Proceedings of Machine Learning Research*, pages 145–156. PMLR, Jul 2017.

[4] Alexander Erreygers and Jasper De Bock.
Computing inferences for large-scale continuous-time Markov chains by combining lumping with imprecision.
In Sébastien Destercke, Thierry Denoeux, María Ángeles Gil, Przemyslaw Grzegorzewski, and Olgierd Hryniewicz, editors, *Uncertainty Modelling in Data Science*, pages 78–86. Springer International Publishing, 2019.

References II

[5] Thomas Krak, Jasper De Bock, and Arno Siebes.
Imprecise continuous-time Markov chains.
International Journal of Approximate Reasoning, 88:452–528, 2017.

[6] Thomas Krak, Alexander Erreygers, and Jasper De Bock.
An imprecise probabilistic estimator for the transition rate matrix of a continuous-time markov chain.
In Sébastien Destercke, Thierry Denoeux, María Ángeles Gil, Przemyslaw Grzegorzewski, and Olgierd Hryniewicz, editors, *Uncertainty Modelling in Data Science*, pages 124–132, 2019.

[7] Marcel F. Neuts.
Matrix-geometric solutions in stochastic models: an algorithmic approach.
Dover, 1981.

[8] Lewis Paton, Matthias C. M. Troffaes, Nigel Boatman, Mohamud Hussein, and Andy Hart.
Multinomial logistic regression on Markov chains for crop rotation modelling.
In Anne Laurent, Oliver Strauss, Bernadette Bouchon-Meunier, and Ronald R. Yager, editors, *Proceedings of the 15th International Conference IPMU 2014 (Information Processing and Management of Uncertainty in Knowledge-Based Systems, 15–19 July 2014, Montpellier, France)*, volume 444 of *Communications in Computer and Information Science*, pages 476–485. Springer, 2014.

References III

[9] Matthias Troffaes, Jacob Gledhill, Damjan Škulj, and Simon Blake.
Using imprecise continuous time Markov chains for assessing the reliability of power networks with common cause failure and non-immediate repair.
In Thomas Augustin, Serena Doria, Enrique Miranda, and Erik Quaeghebeur, editors, *ISIPTA'15: Proceedings of the 9th International Symposium on Imprecise Probability: Theories and Applications*, pages 287–294, Pescara, Italy, July 2015. ARACNE.

[10] Matthias C. M. Troffaes and Simon Blake.
A robust data driven approach to quantifying common-cause failure in power networks.
In F. Cozman, T. Denœux, S. Destercke, and T. Seidenfeld, editors, *ISIPTA'13: Proceedings of the Eighth International Symposium on Imprecise Probability: Theories and Applications*, pages 311–317, Compiègne, France, July 2013. SIPTA.