What if you don’t know your probabilities?

A crash course in imprecise probabilities and their application to Markov chains

Jasper De Bock

25 November 2019
Brighton
What if you don’t know your probabilities?

A crash course in imprecise probabilities and their application to Markov chains.
MODELLING UNCERTAINTY

- p-boxes
- set theory
- logic
- probability measures
- belief functions
- random sets
- imprecise probabilities
- probability intervals
- choice functions
- sets of probability measures
- lower and upper expectations
- sets of desirable gambles
- set theory
- logic
- probability measures
- random sets
Markov chains → imprecise Markov chains
Bayesian networks → credal networks

And much more...

imprecise probabilities

Bounds on probabilities and expectations
Robust (set-valued) decision making
society for
imprecise probabilities
theories and applications
What if you don’t know your probabilities?

A crash course in imprecise probabilities and their application to Markov chains

Jasper De Bock

25 November 2019
Brighton
FLip Foundations Lab on imprecise probabilities

Jasper De Bock

GHENT UNIVERSITY
FLip
Foundations Lab on imprecise probabilities

Gert de Cooman & Jasper De De Bock
What if you don’t know your probabilities?

A crash course in imprecise probabilities and their application to Markov chains

Jasper De Bock
25 November 2019
Brighton
stochastic process

$X_0$  

$X_t$  

$0$  

$t$
Continuous-time stochastic process
Continuous-time stochastic process

\[ P(X_0 = x) = \pi_0(x) \]
\[ P(X_{t+\Delta} = y | X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x) \]
Continuous-time Markov chain

\[ P(X_0 = x) = \pi_0(x) \]
\[ P(X_{t+\Delta} = y | X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x) = P(X_{t+\Delta} = y | X_t = x) \]

Markov assumption

\[ \begin{align*}
X_0 & \quad X_{t_1} & \quad X_{t_n} & \quad X_t & \quad X_{t+\Delta} \\
0 & \quad t_1 & \quad t_n & \quad t & \quad t + \Delta
\end{align*} \]
Continuous-time Markov chain

\[ P(X_0 = x) = \pi_0(x) \]

\[ P(X_{t+\Delta} = y | X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x) \]

\[ = P(X_{t+\Delta} = y | X_t = x) \]

\[ \approx I(x, y) + \Delta Q_t(x, y) \]
Continuous-time Markov chain

\[ P(X_0 = x) = \pi_0(x) \]

\[ P(X_{t+\Delta} = y|X_t = x_t, \ldots, X_{t_n} = x_{t_n}, X_t = x) = P(X_{t+\Delta} = y|X_t = x) \approx I(x, y) + \Delta Q_t(x, y) \]

**assumption of time-homogeneity**
Continuous-time Markov chain

\[ P(X_0 = x) = \pi_0(x) \]  
initial distribution

\[ P(X_{t+\Delta} = y|X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x) \]
\[ = P(X_{t+\Delta} = y|X_t = x) \]
\[ \approx I(x, y) + \Delta Q_t(x, y) \]

\[ \sum_y Q(x, y) = 0 \]
\[ Q(x, x) \leq 0 \]
\[ (\forall y \neq x) Q(x, y) \geq 0 \]

\[
\begin{array}{cccccc}
X_0 & X_{t_1} & X_{t_n} & X_t & X_{t+\Delta} \\
0 & t_1 & t_n & t & t + \Delta
\end{array}
\]
\[ Q = \begin{bmatrix}
-4 & 3 & 1 & 0 \\
4 & -6 & 2 & 0 \\
2 & 3 & -6 & 1 \\
0 & 0 & 2 & -2
\end{bmatrix} \]

**Diagram:**

```
A ---- 2 ----> B
  ^  1  ^
  |    |    |
  v  3  v
  B ---- 4 ----> C
  |    |    |
  v  2  v
  C ---- 2 ----> D
  |    |    |
  v  3  v
  D ---- 1 ----> B
```

**Formula:**

\[ Q(x, y) \]
What is $P(X_t = y|X_0 = x)$?
What is $P(X_t = y | X_0 = x)$?

**Transition matrix:**

$$T_t(x, y) := P(X_t = y | X_0 = x)$$

**Backward Kolmogorov differential equation**

$$\frac{dT_t}{dt} = QT_t, \text{ with } T_0 = I$$

$$\implies T_t = e^{Qt} = \lim_{n \to +\infty} (I + \frac{t}{n}Q)^n$$
What is $P(X_t = y | X_0 = x)$? 

transition matrix: 

$T_t(x, y) := P(X_t = y | X_0 = x)$

backward Kolmogorov differential equation

$$\frac{d}{dt} T_t = QT_t , \text{ with } T_0 = I$$

$$\Rightarrow \quad T_t = e^{Qt} = \lim_{n \to +\infty} (I + \frac{t}{n} Q)^n$$
What is $P(X_t = y | X_0 = x)$? \[ e^{Q_t}(x, y) \]

What is $E(f(X_t) | X_0 = x)$? \[ [e^{Q_t} f](x) \]

$x = 0 \\ f(X_t) = X_t$
What is $P(X_t = y|X_0 = x)$?

$= E(\mathbb{1}_y(X_t)|X_0 = x)$

$\mathbb{1}_y(X_t) = \begin{cases} 1 & \text{if } X_t = y \\ 0 & \text{otherwise} \end{cases}$

What is $E(f(X_t)|X_0 = x)$?

$[e^{Qt}f](x)$
What is $P(X_t = y | X_0 = x)$?

$\pi_\infty(y) := \lim_{t \to +\infty} P(X_t = y | X_0 = x)$

$E_\infty(f) := \lim_{t \to +\infty} E(f(X_t) | X_0 = x)$

What is $E(f(X_t) | X_0 = x)$?

$[e^{Q_t} f](x)$
Reliability engineering (failure probabilities, …)

Queuing theory (waiting in line …)
- optimising supermarket waiting times
- dimensioning of call centers
- airport security lines
- router queues on the internet

Chemical reactions (time-evolution …)

Pagerank

...
So how about imprecision?
So how about imprecision?

What if we don’t know $Q$ exactly?
Imprecise continuous-time Markov chain

\[ P(X_0 = x) = \pi_0(x) \]
\[ P(X_{t+\Delta} = y|X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x) \]
\[ = P(X_{t+\Delta} = y|X_t = x) \]
\[ \approx I(x, y) + \Delta Q_t(x, y) \]
\[ Q \in \mathcal{Q} \]
Imprecise continuous-time Markov chain

\[
P(X_0 = x) = \pi_0(x)
\]

\[
P(X_{t+\Delta} = y|X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x) = P(X_{t+\Delta} = y|X_t = x) 
\]

\[
\approx I(x, y) + \Delta Q_t(x, y)
\]

\[
\overline{E}(f) = \min_{Q \in \mathcal{Q}} E(f)
\]

\[
\underline{E}(f) = \max_{Q \in \mathcal{Q}} E(f) = -\overline{E}(-f)
\]

Q \in \mathcal{Q}

probability bounds are special cases!
Imprecise continuous-time Markov chain

\[
P(X_0 = x) = \pi_0(x)
\]

\[
P(X_{t+\Delta} = y|X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x)
\]

\[
= P(X_{t+\Delta} = y|X_t = x)
\]

\[
\approx I(x, y) + \Delta Q_t(x, y)
\]

assumption of time-homogeneity

\[
\begin{align*}
X_0 & \quad X_{t_1} & \quad X_{t_n} & \quad X_t & \quad X_{t+\Delta} \\
0 & \quad t_1 & \quad t_n & \quad t & \quad t + \Delta
\end{align*}
\]
Imprecise continuous-time Markov chain

\[ P(X_0 = x) = \pi_0(x) \]

\[ P(X_{t+\Delta} = y \mid X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x) \]

\[ = P(X_{t+\Delta} = y \mid X_t = x) \]

\[ \approx I(x, y) + \Delta Q_t(x, y) \]

\[ Q \in Q \]

Assumption of time-homogeneity
Imprecise continuous-time Markov chain

\[ P(X_0 = x) = \pi_0(x) \]

\[ P(X_{t+\Delta} = y|X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x) \]

\[ = P(X_{t+\Delta} = y|X_t = x) \]

\[ \approx I(x, y) + \Delta Q_t(x, y) \]

assumption of

time-homogeneity
Imprecise continuous-time Markov chain

\[ P(X_0 = x) = \pi_0(x) \]
\[ P(X_{t+\Delta} = y | X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x) \]
\[ = P(X_{t+\Delta} = y | X_t = x) \]
\[ \approx I(x, y) + \Delta Q_t(x, y) \]

**Differentiability assumption**
Imprecise continuous-time Markov chain

\[ P(X_0 = x) = \pi_0(x) \]

\[ P(X_{t+\Delta} = y \mid X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x) = P(X_{t+\Delta} = y \mid X_t = x) \approx I(x, y) + \Delta Q_t(x, y) \cap Q \]

- Differentiability assumption
Imprecise continuous-time Markov chain

\[ P(X_0 = x) = \pi_0(x) \]

\[ P(X_{t+\Delta} = y|X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x) = P(X_{t+\Delta} = y|X_t = x) \]

\[ \approx I(x, y) + \Delta Q_{t,\Delta}(x, y) \]

\[ Q \]

\[ X_0 \quad X_{t_1} \quad X_{t_n} \quad X_t \quad X_{t+\Delta} \]

\[ 0 \quad t_1 \quad t_n \quad t \quad t + \Delta \]
Imprecise continuous-time Markov chain

\[ P(X_0 = x) = \pi_0(x) \]

\[ P(X_{t+\Delta} = y | X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x) \]

\[ = P(X_{t+\Delta} = y | X_t = x) \]

\[ \approx I(x, y) + \Delta Q_{t,\Delta}(x, y) \]

Markov assumption

\[
\begin{array}{cccccc}
X_0 & X_{t_1} & X_{t_n} & X_t & X_{t+\Delta} \\
0 & t_1 & t_n & t & t + \Delta \\
\end{array}
\]
Imprecise continuous-time Markov chain

\[ P(X_0 = x) = \pi_0(x) \]
\[ P(X_{t+\Delta} = y|X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x) = P(X_{t+\Delta} = y|X_t = x) \]
\[ \approx I(x, y) + \Delta Q_{t,\Delta} (x, y) \]

Markov assumption
Imprecise continuous-time Markov chain

\[ P(X_0 = x) = \pi_0(x) \]
\[ P(X_{t+\Delta} = y | X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x) \]
\[ \approx I(x, y) + \Delta Q_{t, \Delta, x_1, \ldots, x_n}(x, y) \]

Markov assumption

\[
\begin{align*}
X_0 & \quad X_{t_1} & \quad X_{t_n} & \quad X_t & \quad X_{t+\Delta} \\
0 & \quad t_1 & \quad t_n & \quad t & \quad t + \Delta
\end{align*}
\]
Imprecise continuous-time Markov chain

\[ P(X_0 = x) = \pi_0(x) \]
\[ P(X_{t+\Delta} = y | X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x) \approx I(x, y) + \Delta Q_{t,\Delta,x_1,\ldots,x_n}(x, y) \]

\[ \overline{E}(f) = \min_P \overline{E}(f) \]
\[ \underline{E}(f) = \max_P \underline{E}(f) = -\overline{E}(-f) \]

probability bounds are again just special cases!
What is $E(f(X_t) | X_0 = x)$?
What is $E(f(X_t)|X_0 = x)$?

transition operator: $[T_t(f)](x) = E(f(X_t)|X_0 = x)$

backward Kolmogorov differential equation

$$\frac{d}{dt} T_t = QT_t, \text{ with } T_0 = I$$

$$\implies T_t = e^{Qt} = \lim_{n \to +\infty} (I + \frac{t}{n}Q)^n$$

transition rate operator: $Qf(x) = \min_{Q \in \mathcal{Q}} Qf(x)$
What is $\mathbb{E}(f(X_t)|X_0 = x)$? 

transition operator: $[T_t(f)](x) = \mathbb{E}(f(X_t)|X_0 = x)$

backward Kolmogorov differential equation

$\frac{d}{dt} T_t = QT_t$, with $T_0 = I$

$\Rightarrow T_t = e^{Qt} = \lim_{n \to +\infty} (I + \frac{t}{n}Q)^n$

transition rate operator: $Qf(x) = \min_{Q \in \mathcal{Q}} Qf(x)$
Imprecise continuous-time Markov chain
Imprecise continuous-time Markov chain

\[
\mathbb{E}(f(X_{t+\Delta}|X_t = x) = \mathbb{E}(f(X_{t+\Delta}|X_0 = x)) = [\mathcal{T}_{\Delta}(f)](x)
\]
Imprecise continuous-time Markov chain

\[ E(f(X_{t+\Delta}|X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x) = E(f(X_{t+\Delta}|X_t = x) = E(f(X_{\Delta}|X_0 = x) = [T_\Delta(f)](x) \]

Markov property

Time-homogeneity

\[ E_\infty(f) = \lim_{t \to +\infty} E(f(X_t)|X_0 = x) = \lim_{t \to +\infty} [T_\Delta(f)](x) \]
What if I don’t have any imprecisation?
What if I don’t have any imprecision?

I know $Q$ exactly!
Solving the scaling problem

$X_0$ $X_t$

$0$ $t$

$X_t \in \mathcal{X}$ original Markov chain

$\Lambda: \mathcal{X} \rightarrow \hat{\mathcal{X}}$ lumping map

$\hat{X}_t = \Lambda(X_t) \in \hat{\mathcal{X}}$ lumped process
<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>-4</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>4</td>
<td>-6</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>C</td>
<td>2</td>
<td>3</td>
<td>-6</td>
<td>1</td>
</tr>
<tr>
<td>D</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>-2</td>
</tr>
</tbody>
</table>

\[ Q = \]

**GOOD**
- Amorous

**NOT GOOD**
- Bickering
- Confusion
- Depression
Solving the scaling problem

\[ \hat{Q}_{t, x_{t_1}, \ldots, x_{t_n}} \in \hat{\mathcal{Q}} \]
Solving the scaling problem

\[
E(\hat{f}(\hat{X}_t) | \hat{X}_0 = \hat{x}) \geq \left[ e^{Q_t \hat{f}} \right](\hat{x}) \\
= \left[ \lim_{n \to +\infty} (I + \frac{t}{n} \hat{Q})^n \hat{f} \right](\hat{x})
\]

\[
[\hat{Q} \hat{f}](\hat{x}) = \min \left\{ \sum_{\hat{y} \in \hat{X}} \hat{f}(\hat{y}) \sum_{y \sim \hat{y}} Q(x, y) : x \sim \hat{x} \right\}
\]
Reliability engineering (failure probabilities, …)

Queuing theory (waiting in line …)
- optimising supermarket waiting times
- dimensioning of call centers
- airport security lines
- router queues on the internet

Chemical reactions (time-evolution …)

Pagerank

…
Message passing in optical links

\[ S = m_1 F \]

\[ m_2 = \frac{m_1}{n_2} \] superchannels

\begin{align*}
\text{type I messages require 1 channel} \\
\text{type II messages require 1 superchannel (} n_2 \text{ channels)}
\end{align*}

We want to know the blocking probability of messages for a given policy, and optimise it.
\[ \mathcal{X}_{\text{det}} := \left\{ (i_0, \ldots, i_{n_2}) \in \mathbb{N}^{(n_2+1)} : \sum_{k=0}^{n_2} i_k \leq m_2 \right\} \]

\[ I := \sum_{k=0}^{n_2} i_k \quad R := \sum_{k=0}^{n_2-1} i_k(n_2 - k) \]
\[ \mathcal{K}_{\text{red}} := \{(i, j, e) \in \mathbb{N}^3 : m_2 \leq i + j + e, i + (j + e)n_2 \leq m_1 \} \]

\[ R := m_1 - i - jn_2 \]
(Erreygers et al. 2018)
Advantages of imprecise Markov chains

- Partially specified $Q$ (and $\pi_0$) are allowed
- Time-homogeneity can be relaxed
- The Markov assumption can be relaxed
- Efficient computations remain possible
- State space explosion can be dealt with