A game-theoretic ergodic theorem for imprecise Markov chains

Gert de Cooman

Ghent University, SYSTeMS

gert.decooman@UGent.be
http://users.UGent.be/~gdcooma
gertekoo.wordpress.com

GTP 2014
CIMAT, Guanajuato
13 November 2014
My boon companions

FILIP HERMANS  
ENRIQUE MIRANDA  
JASPER DE BOCK
Jean Ville and martingales
Définition 1. — Soit $X_1, X_2, \ldots, X_n, \ldots$ une suite de variables aléatoires, telle que les probabilités

$$\Pr \left\{ X_1 < x_1, X_2 < x_2, \ldots, X_n < x_n \right\} \quad (n = 1, 2, 3, \ldots)$$

soient bien définies et que les $X_i$ ne puissent prendre que des valeurs finies.

Soit une suite de fonctions $s_0, s_1(x_1), s_2(x_1, x_2), \ldots$ non négatives telles que

$$(14) \left\{ M_{x_1, x_2, \ldots, x_{n-1}} \left\{ s_n(x_1, x_2, \ldots, x_{n-1}, X_n) \right\} = s_{n-1}(x_1, x_2, \ldots, x_{n-1}), \right.$$  

où $M_X\{Y\}$ représente d’une manière générale la valeur moyenne conditionnelle de la variable $Y$ quand on connaît la position du point aléatoire $X$, au sens indiqué par M. P. Lévy.

Dans ces conditions, nous dirons que la suite $\{s_n\}$ définit une martingale ou un jeu équitable.

Étude critique de la notion de collectif, 1939, p. 83
In a (perhaps) more modern notation

Ville’s definition of a martingale

A martingale $s$ is a sequence of real functions $s_o$, $s_1(X_1)$, $s_2(X_1,X_2)$, \ldots such that

1. $s_o = 1$;
2. $s_n(X_1, \ldots, X_n) \geq 0$ for all $n \in \mathbb{N}$;
3. $E(s_{n+1}(x_1, \ldots, x_n, X_{n+1})|x_1, \ldots, x_n) = s_n(x_1, \ldots, x_n)$ for all $n \in \mathbb{N}_0$ and all $x_1, \ldots, x_n$.

It represents the outcome of a fair betting scheme, without borrowing (or bankruptcy).
Ville’s theorem
The collection of all (locally defined!) martingales determines the probability $P$ on the sample space $\Omega$:

\[
P(A) = \sup\{\lambda \in \mathbb{R} : s \text{ martingale and } \limsup_{n \to +\infty} \lambda s_n(X_1, \ldots, X_n) \leq I_A\}
\]

\[
= \inf\{\lambda \in \mathbb{R} : s \text{ martingale and } \liminf_{n \to +\infty} \lambda s_n(X_1, \ldots, X_n) \geq I_A\}
\]
Ville’s theorem
The collection of all (locally defined!) martingales determines the probability $P$ on the sample space $\Omega$:

$$P(A) = \sup\{\lambda \in \mathbb{R} : s \text{ martingale and } \limsup_{n \to +\infty} \lambda s_n(X_1, \ldots, X_n) \leq \mathbb{I}_A\}$$

$$= \inf\{\lambda \in \mathbb{R} : s \text{ martingale and } \liminf_{n \to +\infty} \lambda s_n(X_1, \ldots, X_n) \geq \mathbb{I}_A\}$$

Turning things around
Ville’s theorem suggests that we could take a convex set of martingales as a primitive notion, and probabilities and expectations as derived notions.

That we need an convex set of them, elucidates that martingales are examples of partial probability assessments.
Imprecise probabilities: dealing with partial probability assessments
Partial probability assessments
lower and/or upper bounds for
- the probabilities of a number of events,
- the expectations of a number of random variables
Partial probability assessments
lower and/or upper bounds for
- the probabilities of a number of events,
- the expectations of a number of random variables

Imprecise probability models
A partial assessment generally does not determine a probability measure uniquely, only a convex closed set of them.
Partial probability assessments
lower and/or upper bounds for
  – the probabilities of a number of events,
  – the expectations of a number of random variables

Imprecise probability models
A partial assessment generally does not determine a probability measure uniquely, only a convex closed set of them.

IP Theory
systematic way of dealing with, representing, and making conservative inferences based on partial probability assessments
Lower and upper expectations
A Subject is uncertain about the value that a variable $X$ assumes in $\mathcal{X}$.

Gambles:
A gamble $f : \mathcal{X} \rightarrow \mathbb{R}$ is an uncertain reward whose value is $f(X)$. $\mathcal{G}(\mathcal{X})$ denotes the set of all gambles on $\mathcal{X}$.
Lower and upper expectations

A Subject is uncertain about the value that a variable $X$ assumes in $\mathcal{X}$.

Gambles:
A gamble $f : \mathcal{X} \to \mathbb{R}$ is an uncertain reward whose value is $f(X)$. $\mathcal{G}(\mathcal{X})$ denotes the set of all gambles on $\mathcal{X}$.

Lower and upper expectations:
A lower expectation is a real functional that satisfies:

E1. $\underline{E}(f) \geq \inf f$ [bounds]
E2. $\underline{E}(f + g) \geq \underline{E}(f) + \underline{E}(g)$ [superadditivity]
E3. $\underline{E}(\lambda f) = \lambda \underline{E}(f)$ for all real $\lambda \geq 0$ [non-negative homogeneity]

$\overline{E}(f) := -\overline{E}(-f)$ defines the conjugate upper expectation.
Sub- and supermartingales
An event tree and its situations

Situations are nodes in the event tree, and the sample space $\Omega$ is the set of all terminal situations:
An event $A$ is a subset of the sample space $\Omega$:

$$\Gamma(s) := \{ \omega \in \Omega : s \sqsubseteq \omega \}$$
In each non-terminal situation $s$, **Subject** has a belief model $Q(\cdot |s)$.

$D(s) = \{c_1, c_2\}$ is the set of daughters of $s$. 
We can use the local models $Q(\cdot | s)$ to define sub- and supermartingales:

A submartingale $\underline{M}$
is a real process such that in all non-terminal situations $s$:

$$Q(\underline{M}(s \cdot) | s) \geq \underline{M}(s).$$

A supermartingale $\overline{M}$
is a real process such that in all non-terminal situations $s$:

$$Q(\overline{M}(s \cdot) | s) \leq \overline{M}(s).$$
The most conservative lower and upper expectations on $\mathcal{G}(\Omega)$ that coincide with the local models and satisfy a number of additional continuity criteria (cut conglomerability and cut continuity):

Conditional lower expectations:

$$\underline{E}(f|s) := \sup\{\underline{\mathcal{M}}(s) : \limsup \underline{\mathcal{M}} \leq f \text{ on } \Gamma(s)\}$$

Conditional upper expectations:

$$\overline{E}(f|s) := \inf\{\overline{\mathcal{M}}(s) : \liminf \overline{\mathcal{M}} \geq f \text{ on } \Gamma(s)\}$$
Test supermartingales and strictly null events

A test supermartingale $\mathcal{T}$
is a non-negative supermartingale with $\mathcal{T}(\square) = 1$.
(Very close to Ville’s definition of a martingale.)

An event $A$ is strictly null
if there is some test supermartingale $\mathcal{T}$ that converges to $+\infty$ on $A$:

$$\lim_{n \to \infty} \mathcal{T}(\omega^n) = +\infty \text{ for all } \omega \in A.$$  

If $A$ is strictly null then

$$\overline{P}(A) = \mathcal{E}(\mathbb{I}_A) = \inf\{\overline{\mathbb{M}}(\square) : \lim\inf\overline{\mathbb{M}} \geq \mathbb{I}_A\} = 0.$$
A few basic limit results

**Supermartingale convergence theorem [Shafer and Vovk, 2001]**

A supermartingale $\overline{M}$ that is bounded below converges strictly almost surely to a real number:

$$\liminf \overline{M}(\omega) = \limsup \overline{M}(\omega) \in \mathbb{R} \text{ strictly almost surely.}$$
A few basic limit results

Strong law of large numbers for submartingale differences [De Cooman and De Bock, 2013]
Consider any submartingale $\mathcal{M}$ such that its difference process

$$\Delta \mathcal{M}(s) = \mathcal{M}(s \cdot) - \mathcal{M}(s) \in \mathcal{G}(D(s)) \quad \text{for all non-terminal } s$$

is uniformly bounded. Then $\liminf \langle \mathcal{M} \rangle \geq 0$ strictly almost surely, where

$$\langle \mathcal{M} \rangle(\omega^n) = \frac{1}{n} \mathcal{M}(\omega^n) \quad \text{for all } \omega \in \Omega \text{ and } n \in \mathbb{N}$$
A few basic limit results

Lévy’s zero–one law [Shafer, Vovk and Takemura, 2012]
For any bounded real gamble $f$ on $\Omega$:

$$\limsup_{n \to +\infty} E(f|\omega^n) \leq f(\omega) \leq \liminf_{n \to +\infty} E(f|\omega^n)$$ strictly almost surely.
Imprecise Markov chains
A simple discrete-time finite-state stochastic process

\[ Q(\cdot | \square) \]

- \( Q(\cdot | a) \)
  - \( (a, a) \) \( Q(\cdot | a, a) \) \( (a, a, a) \)
  - \( (a, b) \) \( Q(\cdot | a, b) \) \( (a, b, a) \)
  - \( (a, b) \) \( Q(\cdot | a, b) \) \( (a, b, b) \)

- \( Q(\cdot | b) \)
  - \( (b, b) \) \( Q(\cdot | b, b) \) \( (b, b, b) \)
  - \( (b, a) \) \( Q(\cdot | b, a) \) \( (b, a, a) \)
  - \( (b, a) \) \( Q(\cdot | b, a) \) \( (b, a, b) \)
  - \( (b, b) \) \( Q(\cdot | b, b) \) \( (b, b, b) \)
An imprecise IID model
An imprecise Markov chain

\[
\begin{align*}
Q(\cdot | □) &
\begin{cases}
(b, b) & Q(\cdot | b) \\
(b, a) & Q(\cdot | a)
\end{cases} \\
&
\begin{cases}
(b, b) & Q(\cdot | b) \\
(a, b) & Q(\cdot | b)
\end{cases} \\
&
\begin{cases}
(b, b, b) & (b, a, b) & (a, b, b) \\
(b, b, a) & (b, a, a) & (a, b, a) \\
(b, b, b) & (b, a, b) & (a, b, a) \\
(b, a, a) & (a, a, b) & (a, a, a)
\end{cases}
\end{align*}
\]
Stationarity and ergodicity

The lower expectation $E_n$ for the state $X_n$ at time $n$:

$$E_n(f) = E(f(X_n))$$

The imprecise Markov chain is Perron–Frobenius-like if for all marginal models $E_1$ and all $f$:

$$E_n(f) \to E_\infty(f).$$

and if $E_1 = E_\infty$ then $E_n = E_\infty$, and the imprecise Markov chain is stationary.

In any Perron–Frobenius-like imprecise Markov chain:

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{k=1}^{n} E_n(f) = E_\infty(f)$$

and

$$E_\infty(f) \leq \liminf_{n \to +\infty} \frac{1}{n} \sum_{k=1}^{n} f(X_k) \leq \limsup_{n \to +\infty} \frac{1}{n} \sum_{k=1}^{n} f(X_k) \leq \overline{E}_\infty(f) \text{ str. almost surely.}$$
Introduce a shift operator:

$$\theta \omega = \theta(x_1, x_2, x_3, \ldots) := (x_2, x_3, x_4, \ldots)$$

for all $$\omega \in \Omega$$,

and for any gamble $$f$$ on $$\Omega$$ a shifted gamble $$\theta f := f \circ \theta$$:

$$(\theta f)(\omega) := f(\theta \omega)$$

for all $$\omega \in \Omega$$.

For any bounded gamble $$f$$ on $$\Omega$$, the bounded gambles:

$$g = \liminf_{n \to +\infty} \frac{1}{n} \sum_{k=0}^{n-1} \theta^k f$$

and

$$g = \limsup_{n \to +\infty} \frac{1}{n} \sum_{k=0}^{n-1} \theta^k f$$

are shift-invariant: $$\theta g = g$$. 
A more general ergodic theorem: use Lévy’s zero–one law

In any Perron–Frobenius-like imprecise Markov chain, for any shift-invariant gamble \( g = \theta g \) on \( \Omega \):

\[
\lim_{n \to +\infty} E(g \mid \omega^n) = E_\infty(g) \quad \text{and} \quad \lim_{n \to +\infty} \overline{E}(g \mid \omega^n) = \overline{E}_\infty(g)
\]

and therefore

\[
E_\infty(g) \leq g \leq \overline{E}_\infty(g) \quad \text{strictly almost surely.}
\]
New books