Finite Additivity in Dubins-Savage Gambling and Stochastic Games

Bill Sudderth
University of Minnesota

This talk is based on joint work with Lester Dubins, David Heath, Ashok Maitra, and Roger Purves.
References

I will survey some measure theoretic aspects of 50 years of research in Dubins-Savage gambling and its connection to stochastic games. Most references are left to the end. The basic reference is of course:

A Dubins-Savage gambling problem: $(S, \Gamma, g)$

$S$ - the **state space**

$\Gamma$ - the **gambling house** - for each $x \in S$, $\Gamma(x)$ is a nonempty set of possible distributions for the next state.

A player starts at some $x_0 \in S$ and chooses $\sigma_0 \in \Gamma(x_0)$. The next state $X_1$ has distribution $\sigma_0$. Given $X_1 = x_1$, the player chooses $\sigma_1(x_1) \in \Gamma(x_1)$ as the conditional distribution of $X_2$. Given $X_1 = x_1, X_2 = x_2$, the player choses $\sigma_2(x_1, x_2) \in \Gamma(x_2)$ as the conditional distribution of $X_3$. And so on.

The sequence $\sigma = (\sigma_0, \sigma_1, \ldots)$ is a **strategy**.

$g : S^\infty \mapsto \mathbb{R}$ - the **payoff function**. The player seeks to choose $\sigma$ to maximize the expected value $E_{\sigma}g(x_1, x_2, \ldots)$.
A finitely additive formulation

$S$ is an arbitrary nonempty set.

For each $x \in S$, $\Gamma(x)$ is a nonempty set of gambles on $S$. (A gamble on a set $S$ is a finitely additive probability measure defined on all subsets of $S$.)

The payoff function $g : S^\infty \mapsto \mathbb{R}$ is bounded and $\mathcal{F}$-measurable where $\mathcal{F}$ is the sigma-field generated by the open subsets of $S^\infty$ when $S$ has the discrete topology and $S^\infty$ has the product topology.

The definition of $E_{\sigma}g$ is in three steps.
Step 1. The definition of $E_\sigma g = \int g \, dP_\sigma$ for a space of functions $g : S^\infty \mapsto \mathbb{R}$

Dubins and Savage recast the conditioning formula $E[Y] = E[E[Y|X]]$ as

$$E_\sigma g = \int E_\sigma[x_1](gx_1) \, \sigma_0(dx_1).$$

where $\sigma[x_1]$ is the conditional strategy that follows $\sigma$ after the first stage and $gx_1$ is the section of $g$ such that $(gx_1)(x_2, x_3, \ldots) = g(x_1, x_2, x_3, \ldots)$.

This formula together with the requirement that $E_\sigma c = c$ for constants $c$ determines $E_\sigma$ on a linear space of functions $g$ that includes the indicators of all clopen subsets of $S^\infty$ when $S$ has the discrete topology and $S^\infty$ has the product topology.
Step 2. The Lebesgue-like extension to open sets

For open subsets $O \subseteq S^\infty$, let

$$P_\sigma(O) = \sup\{P_\sigma(K) : K \subseteq O, K \text{ clopen}\}.$$  

Dubins (1974) showed $P_\sigma$ is then finitely additive on the lattice of open sets with a unique extension, also written as $P_\sigma$, to the algebra generated by the open sets.

Non-uniqueness: There may be other finitely additive extensions to the open sets, but this one has many good properties.
Step 3. A further extension by squeezing

Let $\mathcal{A}$ be the collection of all $A \subseteq S^\infty$ such that
\[
\inf \{ P_\sigma(O) : A \subseteq O, \text{ } O \text{ open} \} = \sup \{ P_\sigma(C) : C \subseteq A, \text{ } C \text{ closed} \}.
\]
Write $P_\sigma(A)$ for this common value.

Then $P_\sigma$ is finitely additive on $\mathcal{A}$ and $\mathcal{A}$ contains the sigma-field $\mathcal{F}$ generated by the open sets. The sigma-field $\mathcal{F}$ includes all the sets usually considered in countably additive probability.

The expectation $E_\sigma g$ is well-defined for all bounded, $\mathcal{F}$-measurable functions $g : S^\infty \rightarrow \mathbb{R}$.

The classical limit theorems such as the strong law of large numbers and the martingale convergence theorem generalize to this finitely additive setting (Karandikar, 1982, 1988).
Examples of payoff functions

Let \( r : S \to \mathbb{R} \) be a bounded daily reward function.

**Discounted:** \( g(x_1, x_2, \ldots) = \sum_{n=1}^{\infty} \beta^{n-1} r(x_n) \) where \( 0 < \beta < 1 \).

**Long run average:** \( g(x_1, x_2, \ldots) = \lim \sup_m \left( \frac{1}{m} \sum_{n=1}^{m} r(x_n) \right) \).

These are special cases of:

**Limsup:** \( u^*(x_1, x_2, \ldots) = \lim \sup_n u(x_1, \ldots, x_n) \)

Here \( u \) is a bounded real-valued function defined for all finite sequences \( (x_1, \ldots, x_n) \) of states.
The optimal reward function $\Gamma_{g}$

For $x \in S$,

$$\Gamma_{g}(x) = \sup E_{\sigma}[g]$$

where the sup is over all strategies $\sigma$ available at $x$ in $\Gamma$.

Main problems: Calculate $\Gamma_{g}$ and find optimal or nearly optimal strategies.
A regularity property of the optimal reward

Let \( B \in \mathcal{F} \) and \( \sigma \) be a strategy. Then, by the definition of \( P_\sigma \),

\[
P_\sigma(B) = \inf \{ P_\sigma(O) \mid B \subseteq O, \ O \text{ open} \}.
\]

The optimal reward has the same property.

Write \( \Gamma B(x) = \Gamma 1_B(x) = \sup \{ P_\sigma(B) \mid \sigma \text{ at } x \} \).

**Theorem 1.** For \( x \in S \), \( \Gamma B(x) = \inf \{ \Gamma O(x) \mid B \subseteq O, \ O \text{ open} \} \),

The equality can be rewritten as:

\[
\sup_{\sigma \text{ at } x} \inf_{O \supseteq B} P_\sigma(O) = \inf_{O \supseteq B} \sup_{\sigma \text{ at } x} P_\sigma(O)
\]
The one-day operator $T$

For bounded functions $f : S \mapsto \mathbb{R}$ and $x \in S$, define

$$Tf(x) = \sup \{ \int f(x_1) \gamma(dx_1) | \gamma \in \Gamma(x) \}$$

For an open set $O$, the optimal reward function $\Gamma O$ can be calculated by a transfinite form of backward induction. The algorithm for $\Gamma O$ is completely determined by $T$. So, by the regularity property, the value of $\Gamma B$ is also determined by $T$ for all $B \in \mathcal{F}$. 
An example. Let $E \subseteq S$ and let $O = \bigcup_n [x_n \in E]$. Assume that $\delta(x) \in \Gamma(x)$ for every $x \in S$.

Define functions $U_\xi : S \mapsto \mathbb{R}$ for ordinals $\xi$ by setting $U_0 = 1_E$ and, for $\xi > 0$,

$$U_\xi = \begin{cases} TU_{\xi-1}, & \text{if } \xi \text{ is a successor} \\ \sup_{\eta < \xi} U_\eta, & \text{if not.} \end{cases}$$

Then $\Gamma O = U_{\xi^*}$ where $\xi^*$ is the least ordinal such that $TU_{\xi^*} = U_{\xi^*}$.

For $S$ finite, $\xi^* = \omega$ and $\Gamma O = \lim_n U_n$. 
Another regularity property

Let $g : S^\infty \mapsto \mathbb{R}$ be bounded and $\mathcal{F}$-measurable.

**Theorem 2.** For $x \in S$, $\Gamma g(x) = \sup\{\Gamma u^*(x) | u^* \leq g\}$.

Here $u^*(x_1, x_2, \ldots) = \limsup_n u(x_1, \ldots, x_n)$ is the limsup payoff and $u$ ranges over all bounded, real-valued $u$.

There is a transfinite algorithm for calculating the functions $\Gamma u^*$. The algorithm is completely determined by the one-day operator $T$. So $\Gamma g$ is also determined by $T$. 
A measurable, countably additive formulation (Strauch, 1967)

1. The state space $S$ is a nonempty Borel subset of a Polish space.

2. For every $x \in S$, $\Gamma_M(x)$ is a nonempty set of countably additive probability measures defined on the sigma-field $B(S)$ of Borel subsets of $S$.

3. $\Gamma_M(x)$ varies “measurably in $x$” in the sense that the set $\{(x, \gamma) | \gamma \in \Gamma_M(x)\}$ is a Borel subset of $S \times C(S)$ where $C(S)$ is the set of countably additive probability measures on $B(S)$ with its usual topology.

4. The payoff function $g : S^\infty \mapsto \mathbb{R}$ is bounded and Borel measurable when $S^\infty$ is given its product Borel sigma-field.
Measurable strategies

A strategy $\sigma = (\sigma_0, \sigma_1, \ldots)$ is measurable and available at $x$ in $\Gamma_M$ if $\sigma_0 \in \Gamma_M(x)$ and, for all $n \geq 1$ and all $(x_1, \ldots, x_n) \in S^n$, $\sigma_n(x_1, \ldots, x_n) \in \Gamma_M(x_n)$ and $\sigma_n$ is a universally measurable function from $S^n$ to $C(S)$.

A measurable strategy $\sigma$ determines a countably additive probability measure $P^M_\sigma$ on the Borel subsets of $S^\infty$ with expectation operator $E^M_\sigma$.

The optimal reward for a measurable problem with payoff function $g$ is, for $x \in S$,

$$Mg(x) = \sup E^M_\sigma g$$

where the sup is over all measurable $\sigma$ at $x$. The function $Mg$ need not be Borel, but is universally measurable.
A regularity property of $M$

Let $g : S^\infty \mapsto \mathbb{R}$ be bounded and Borel measurable.

**Theorem 3.** For $x \in S$, $Mg(x) = \sup\{Mu^*(x) | u^* \leq g\}$, where $u^*(x_1, x_2, \ldots) = \lim \sup_n u(x_1, \ldots, x_n)$ and $u$ ranges over all bounded, real-valued, Borel measurable $u$.

As in the finitely additive case, there is a transfinite algorithm for calculating the functions $Mu^*$.
Finitely additive extensions of measurable problems

Every countably additive measure $\gamma_M$ available in a measurable house $\Gamma_M$ can be extended to a finitely additive additive measure $\gamma$ defined on all subsets of $S$. (The measure $\gamma$ is typically not unique.)

Suppose every $\gamma_M$ is extended in this way and let $\Gamma$ be the resulting finitely additive house.

**Question:** How are the optimal reward functions $Mg$ and $\Gamma g$ related?
Theorem 4. Suppose $\Gamma_M$ is a measurable gambling house defined on the Borel state space $S$ and that $\Gamma$ is a finitely additive extension of $\Gamma_M$. If $g : S^\infty \mapsto \mathbb{R}$ is bounded and Borel measurable, then $\Gamma g = Mg$.

So finitely additive gambling theory can be viewed as a generalization of the measurable, countably additive theory.

Also a player in a measurable house cannot improve his payoff by using a nonmeasurable strategy.

The proof of Theorem 4 uses the regularity properties of the operators $\Gamma$ and $M$. 
A two-person, zero-sum game: $G = (A, B, f)$.

$A, B$ - nonempty sets of actions for players 1 and 2.

$f : A \times B \mapsto \mathbb{R}$ - a bounded, real-valued payoff function.

Player 1 chooses an action $a \in A$ and simultaneously player 2 chooses $b \in B$; player 2 pays player 1 the amount $f(a, b)$. The players may choose their actions independently and at random using probability distributions $\mu$ and $\nu$ on $A$ and $B$.

Player 1 seeks to maximize and player 2 to minimize the expected payoff $E_{\mu, \nu} f$. 
The game $G = (A, B, f)$ has a **value** if

$$\sup_{\mu} \inf_{\nu} E_{\mu,\nu} f = \inf_{\nu} \sup_{\mu} E_{\mu,\nu} f.$$ 

**Theorem 5.** (von Neumann, 1928) If $A$ and $B$ are finite, then the game has a value.

If $A$ and $B$ are countably infinite, and probability distributions are required to be countably additive, then the game may not have a value.

**Example.** (Wald) Let $A = B = \{1, 2, \ldots\}$ and let $f(a, b) = 1$ or $-1$ according as $a \geq b$ or $a < b$. 
Games with finitely additive mixed actions

Recall that a gamble on a set is a finitely additive probability measure defined on all subsets.

**Theorem 6.** Suppose
1. $A$ and $B$ are nonempty sets,
2. the distributions $\mu$ and $\nu$ vary over all gambles on $A$ and $B$, respectively,
3. $f : A \times B \to \mathbb{R}$ is a bounded function,
4. the expectation $E_{\mu,\nu} f = \int_B \int_A f(a,b) \mu(da) \nu(db)$ is a double integral in a given order.

Then the game $(A, B, f)$ has a value.

See Marinacci (1997) and Flesch, Vermeulen, and Zseleva (2017) for more on finitely additive values.
A coherence lemma. Let $X$ be the Banach space of all bounded functions $g : B \mapsto \mathbb{R}$ with the supremum norm. Suppose $F$ is a convex subset of $X$ such that $\inf g \leq 0$ for all $g \in F$. (There is no sure win in $F$.) Then there is a gamble $\nu$ on $B$ such that $\int g \, d\nu \leq 0$ for all $g \in F$.

Proof. Let $N = \{ g \in X | \inf g > 0 \}$. Then $N$ is convex with a nonempty interior and $N \cap F = \emptyset$. By a separation theorem, there exists a linear functional $x \in X^*$ such that $xg \leq 0$, for $g \in F$ and $xg \geq 0$ for $g \in N$. Normalize $x$ so that $x(1) = 1$ and take $\nu = x$. 
Proof of Theorem 6. To show
\[
\sup_{\mu} \inf_{\nu} \int \int f(a, b) \mu(da) \nu(db) = \inf_{\nu} \sup_{\mu} \int \int f(a, b) \mu(da) \nu(db),
\]
note that \( \leq \) is always true. To prove \( \geq \), suppose
\( \sup_{\mu} \inf_{\nu} E_{\mu, \nu} f < r \) for some \( r \in \mathbb{R} \). Then
\[
(\forall \mu)(\exists \nu) \left( \int \int f(a, b) \mu(da) \nu(db) < r \right).
\]
So
\[
(\forall \mu)(\exists b) \left( \int f(a, b) \mu(da) < r \right).
\]
Thus every function \( g_{\mu}(b) = \int f(a, b) \mu(da) - r \) has \( \inf_{b} g_{\mu}(b) \leq 0 \). By the lemma, there exists \( \nu^{*} \) such that
\[
\int \int f(a, b) \mu(da) \nu^{*}(db) - r = \int g_{\mu} \, d\nu^{*} \leq 0
\]
for all \( \mu \). So
\[
\inf_{\nu} \sup_{\mu} \int \int f(a, b) \mu(da) \nu(db) \leq \sup_{\mu} \left[ \int \int f(a, b) \mu(da) \nu^{*}(db) \right] \leq r.
\]
Finitely Additive Nash Equilibria (NE)

Two-person, nonzero-sum Game: \( G = (A, B, f_1, f_2) \).

Payoffs: \( E_{\mu,\nu}f_i = \int_A \int_B f_i(a, b)\mu(da)\nu(db) \) to player \( i \) for \( i = 1, 2 \).

NE: \( \mu^*, \nu^* \) where \( \mu^* \) is optimal for player 1 vs \( \nu^* \) and \( \nu^* \) is optimal for player 2 vs \( \mu^* \).

An example of Flesch and Predtetchinski (2018) shows a NE need not exist. Whether there always exist \( \epsilon \)-NE is an interesting question.
Two-Person Zero-Sum Stochastic Game - Finitely Additive

Six ingredients: \( S, A, B, p, g \).

\( S \) - state space

\( A, B \) - action sets for players 1 and 2

\( p(\cdot|x, a, b) \) - law of motion - a gamble on \( S \) for every \( (x, a, b) \in S \times A \times B \)

\( g \) - payoff function from player 2 to player 1 - a bounded function from \((A \times B \times S)\infty\) to \( \mathbb{R} \) that is measurable for the sigma-field \( \mathcal{G} \) generated by the open subsets of \((A \times B \times S)\infty\) when \( A, B, S \) have their discrete topologies and the product has the product topology
Play of the game

Play begins at some state $x_0 \in S$, player 1 chooses $a_0 \in A$ using a gamble on $A$, player 2 chooses $b_0 \in B$ using a gamble on $B$. The next state $x_1$ has distribution $p(\cdot|x_0, a_0, b_0)$ and play continues from $x_1$.

At the next stage, player 1 chooses $a_1 \in A$ using a gamble on $A$, player 2 chooses $b_1 \in B$ using a gamble on $B$. The next state $x_2$ has distribution $p(\cdot|x_1, a_1, b_1)$.

As play continues the players generate an infinite sequence

$$(a_0, b_0, x_1, a_1, b_1, x_2, \ldots) \in (A \times B \times S)^\infty$$

and player 2 pays player 1 the amount

$$g(a_0, b_0, x_1, a_1, b_1, x_2, \ldots).$$
Strategies for the players

Let $Z = A \times B \times S$. A strategy $\pi$ for player 1 is a sequence $\pi_0, \pi_1, \ldots$ such that $\pi_0$ is a gamble on $A$, and for $n \geq 1$ and $(z_1, \ldots, z_n) \in Z^n$, $\pi_n(z_1, \ldots, z_n)$ is a gamble on $A$. A strategy $\rho$ for player 2 is defined similarly.

Strategies $\pi$ and $\rho$ together with the law of motion $p$ determine a Dubins-Savage strategy $\sigma = \sigma_0, \sigma_1 \ldots$ on $Z^\infty$. To define $\sigma$ associate to $x \in S$ and gambles $\mu$ on $A$ and $\nu$ on $B$ the gamble $m = m(x, \mu, \nu)$ on $Z$ where

$$m(d(a, b, x_1))) = p(dx_1|x, a, b)\mu(da)\nu(db).$$

For the initial state $x_0$, let

$$\sigma_0 = m(x_0, \pi_0, \rho_0), \quad \sigma_n(z_1, \ldots, z_n) = m(x_n, \sigma_n(z_1, \ldots, z_n), \rho_n(z_1, \ldots, z_n)).$$
Existence of the value

Let $S, A, B, p, g$ be a finitely additive stochastic game. Write $E_{x,\pi,\rho}[g]$ for the expected reward determined by an initial state $x$ and strategies $\pi$ and $\rho$. The game has a value at state $x$ if

$$\sup_{\pi} \inf_{\rho} E_{x,\pi,\rho}[g] = \inf_{\rho} \sup_{\pi} E_{x,\pi,\rho}[g].$$

This quantity is then the value and is written $V_g(x)$.

**Theorem 7.** If the payoff function $g$ is bounded and $\mathcal{G}$-measurable, then the finitely additive stochastic game has a value $V_g(x)$ for every initial state $x$.

This cannot be true for a general countably additive setting even if the payoff function $g$ depends only on $a_0$ and $b_0$. This follows from Wald’s example.
A regularity property of $V$

Let $g : Z^{\infty} \mapsto \mathbb{R}$ be bounded and $G$-measurable.

**Theorem 8.** For $x \in S$, $Vg(x) = \sup\{Vu^{*}(x) | u^{*} \leq g\}$, where $u^{*}(z_{1}, z_{2}, \ldots) = \lim \sup_{n} u(z_{1}, \ldots, z_{n})$ and $u$ ranges over all bounded, real-valued, $G$-measurable $u$.

Also there is a transfinite algorithm for calculating the value functions $Vu^{*}$. 
Martin’s Theorem: “The determinacy of Blackwell Games”

D. A. Martin (1998) proved Theorem 7 in a countably additive setting with a countable state space $S$ and finite action sets $A$ and $B$. The proof of Theorem 7 adapts Martin’s proof to the general finitely additive setting.

Measurability problems arise if we try to adapt Martin’s proof to a Borel measurable setting.

**Question.** Is there a countably additive, Borel measurable analogue of Theorem 7?
Measurable Stochastic Games

1. $S, A, B$ are nonempty Borel subsets of Polish spaces.

2. $B$ is compact.

3. $p(\cdot|x, a, b)$ is a regular conditional distribution on $S$ given $S \times A \times B$.

4. For every Borel subset $E$ of $S$ and $(x, a) \in S \times A$, $p(E|x, a, \cdot)$ is continuous on $B$.

5. The payoff function $g : S \mapsto \mathbb{R}$ is a bounded Borel measurable function of $(x_1, x_2, \ldots)$. It does not depend on the actions.
A theorem and a possibility

**Theorem 9.** A measurable stochastic game with a limsup payoff \( u^*(x_1, x_2, \ldots) = \limsup_n u(x_1, x_2, \ldots, x_n) \) has a value \( V u^*(x) \) for each initial state \( x \), if \( u \) is a bounded Borel measurable function defined for all finite sequences of states. Also, there is a transfinite algorithm for calculating the value.

**Possibility.** Perhaps every measurable stochastic game satisfying conditions 1 through 5 has a value \( V g(x) \) for each initial state \( x \), and \( V g(x) = \sup \{ V u^*(x) \mid u^* \leq g \} \).
Some references


   itability theorem in finitely additive gambling. *Rendiconti di
   Matematica e delle sue applicazioni* 11 819-842.

    pacitability theorem in measurable gambling theory. *Trans.
    Amer. Math. Soc.* 333 221-249.


    Stochastic Games*. Springer.


