Expressive Rationality

Larry Blume

Cornell University & The Santa Fe Institute & IHS, Vienna

What Caused the Financial Crisis?

What Caused the Financial Crisis?

$$\Pr[T_A<1, T_B<1] = \phi_2(\phi^{-1}(F_A(1)), \phi^{-1}(F_B(1)), \gamma)$$

Here's what killed your 401(k) *David X. Li's Gaussian copula function as first published in 2000. Investors exploited it as a quick—and fatally flawed—way to assess risk. A shorter version appears on this month's cover of* Wired.

Probability
Specifically, this is a joint default probability—the likelihood that any two members of the pool (A and B) will both default. It's what investors are looking for, and the rest of the formula provides the answer.

Survival times
The amount of time between now and when A and B can be expected to default. Li took the idea from a concept in actuarial science that charts what happens to someone's life expectancy when their spouse dies.

Equality
A dangerously precise concept, since it leaves no room for error. Clean equations help both quants and their managers forget that the real world contains a surprising amount of uncertainty, fuzziness, and precariousness.

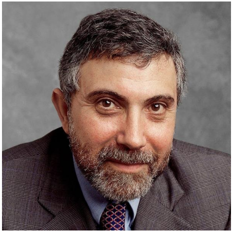
Copula
This couples (hence the Latinate term copula) the individual probabilities associated with A and B to come up with a single number. Errors here massively increase the risk of the whole equation blowing up.

Distribution functions
The probabilities of how long A and B are likely to survive. Since these are not certainties, they can be dangerous: Small miscalculations may leave you facing much more risk than the formula indicates.

Gamma
The all-powerful correlation parameter, which reduces correlation to a single constant —something that should be highly improbable, if not impossible. This is the magic number that made Li's copula function irresistible.

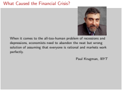
MATH!

What Caused the Financial Crisis?



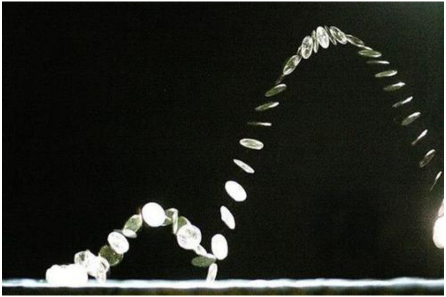
When it comes to the all-too-human problem of recessions and depressions, economists need to abandon the neat but wrong solution of assuming that everyone is rational and markets work perfectly.

Paul Krugman, *NYT*



Copula image source:
http://www.wired.com/techbiz/it/magazine/17-03/wp_quant.

What is Probability?



Coin Flip

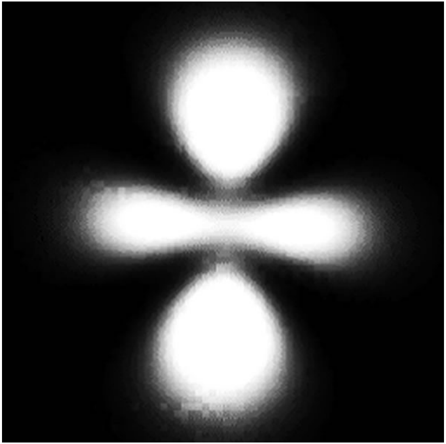
What is Probability?

Horse Race

Horse Race

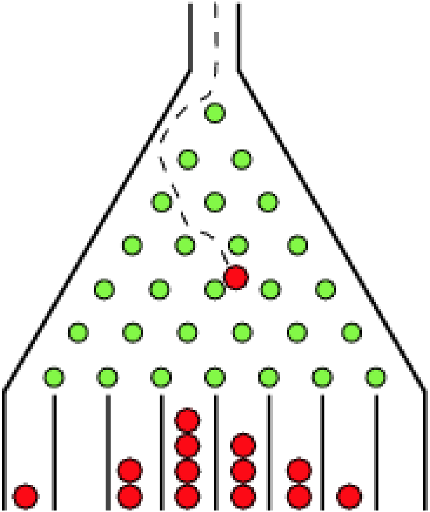
This is the Objective Subjective Distinction

What is Probability?



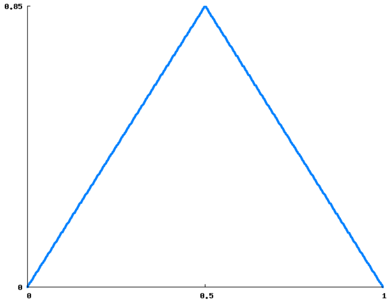
Electron Cloud

What is Probability?



Galton Board

What is Probability?



$$f(x) = \begin{cases} \mu x & \text{if } 0 \leq x \leq 1/2, \\ \mu(1-x) & \text{if } 1/2 < x \leq 1. \end{cases}$$

Tent Map

What is Probability?

Financial Market

Financial Market

The ontic/aleatory — epistemic distinction.

Quantum mechanics gives a probalistic description of the location of an electron around the nucleus of an atom. It is inherently probalistic.

The Dalton board would be deterministic if we knew the initial conditions, but we don't. Same with Coin flips.

For $\sqrt{2} \leq \mu \leq 2$ the Tent map has a unique invariant measure ν_μ which is absolutely continuous with respect to Lebesgue measure, and that measure is ergodic. If you do not know the initial condition, and have any prior belief p_0 on x_0, then your prior belief about x_{1000} should be $p_{1000} \approx \nu_\mu$. So even though beliefs about one run of the tent map should be thought of as epistemic, the system forces your beliefs about x_{1000}. There are physical systems with this property — the location of a gas molecule, solutions to the n-body problem, etc.

Financial Markets?

Classification of Probability Types

	Objective	**Subjective**
Ontic	Electron Cloud	Tent Map
Epistemic	Coin Flip	Horse Race

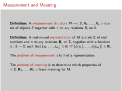
Measurement and Meaning

Definition: A measurement structure $M = < S, R_1, \ldots, R_n >$ is a set of objects S together with n m_i-ary relations R_i on S.

Definition: A real-valued representation of M is a set Σ of real numbers and n m_i-ary relations $\mathbf{R}_i$ on Σ, together with a function $\phi : S \rightarrow \Sigma$ such that $(s_1, \ldots, s_{m_i}) \in R_i$ iff $\big(\phi(s_1), \ldots, \phi(s_{m_i})\big) \in \mathbf{R}_i$.

The problem of measurement is to find a representation.

The problem of meaning is to determine which properties of $< \Sigma, \mathbf{R}_1, \ldots, \mathbf{R}_n >$ have meaning for M.



Measurement and Meaning

Definition: A measurement structure $M =< S, R_1, \ldots, R_n >$ is a set of objects S together with n m_i-ary relations R_i on S.

Definition: A real-valued representation of M is a set Σ of real numbers and n m_i-ary relations $\mathbf{R}_i$ on Σ, together with a function $\phi : S \to \Sigma$ such that $(s_1, \ldots, s_{m_i}) \in R_i$ iff $(\phi(s_1), \ldots, \phi(s_{m_i})) \in \mathbf{R}_i$.

The problem of measurement is to find a representation.

The problem of meaning is to determine which properties of $< \Sigma, \mathbf{R}_1, \ldots, \mathbf{R}_n >$ have meaning for M.

What does it mean to measure something? What are the experiments.

- The relation being measured: warmer than, better than, more likely than; together with operations on objects, such as concatenation, piling up on one side of a balance beam, . . .
- These relations collectively define experiments.

Frequentism

1. Probability is assigned only to collectives. These are repetitive events or mass phenomena.
2. Collectives are modeled as infinite sequences.
3. Relative frequencies converge not just for the infinite sequence, but for any infinite subsequence.

So what is the experiment that we can perform on, say, financial markets?

Probability as a Theory of Measurement

- An algebra $\mathcal{A}$ of sets of elements of a ground set S (of states).
- A complete relation $\succeq B$ on $\mathcal{A}$: $A \succeq B$ means "A is at least as likely as B."
- $S \succeq A \succeq \emptyset$ for all $A \in \mathcal{A}$, and $S \succ \emptyset$.

A *representation of* $\succeq$ is a function $p : \mathcal{A} \to [0, 1]$ such that

$$A \succeq B \text{ iff } p(A) \geq p(B).$$

Probability as a Theory of Measurement: Finite States

When is a representation a probability?

$$A \cap C = \emptyset \quad \Rightarrow \quad p(A \cup C) = p(A) + p(C).$$

Disjoint Union Property:

Suppose C is disjoint from A, B. $A \succeq B$ iff $A \cup C \succeq B \cup C$.

Probability as a Theory of Measurement: Finite States

When is a representation a probability?

$$A \cap C = \emptyset \quad \Rightarrow \quad p(A \cup C) = p(A) + p(C).$$

Cancellation:

If $A_1, \ldots, A_N$ and $B_1, \ldots, B_N$ are sets such that for all s,

$$\#\{n : s \in A_n\} = \#\{n : s \in B_n\}$$

and for all $n \leq N - 1$, $A_n \succeq B_n$, then $B_N \succeq A_N$.

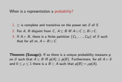
Probability as a Theory of Measurement: Many States

When is a representation a probability?

1. $\succeq$ is complete and transitive on the power set $\mathcal{S}$ of S.
2. For A, B disjoint from C, $A \succeq B$ iff $A \cup C \succeq B \cup C$.
3. If $A \succ B$, there is a finite partition $\{C_1, \ldots, C_M\}$ of S such that for all m, $A \succ B \cup C$.

Theorem (Savage): If so there is a unique probability measure p on $\mathcal{S}$ such that $A \succeq B$ iff $p(A) \geq p(B)$. Furthermore, for all $A \succ \emptyset$ and $0 \leq \rho \leq 1$ there is a $B \subset A$ such that $p(B) = \rho p(A)$.

Probability as a Theory of Measurement: Many States



- If you accept the continuum hypothesis, then p must be only finitely additive.
- $\mathcal{S}$ can be a σ-algebra, but not just an algebra.
- Savage's axioms does not imply that the state space be uncountable.

Alternative Measures of Probability

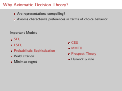
- Sets of Probabilities
- Non-additive Probabilities
- Belief Functions
- Inner and Outer Measure
- Lexicographic Probabilities
- Possibility Measures
- Plausibility Measures
- Ranking Functions

Why Axiomatic Decision Theory?

- Are representations compelling?
- Axioms characterize preferences in terms of choice behavior.

Important Models

- SEU
- LSEU
- Probabilistic Sophistication
- Wald citerion
- Minimax regret
- CEU
- MMEU
- Prospect Theory
- Hurwicz α rule



Representations per se are not compelling.

- Characterize representations in terms of choice behavior.
- A decision model is normatively appropriate iff its characterizing axioms have normative appeal
- A decision model is descriptively appropriate iff its characterizing axioms have descriptive appeal
- Axioms give us a handle on verifying or falsifying, justifying or criticizing given models.

- vN-M - preferences over probability distributions
- SEU

A Common Framework

States:	A finite state space S.
Outcomes:	A finite set O, with best and worst outcomes x^* and x_*.
Roulette Wheels:	The set $\mathbf{R}$ of probability distributions on O.
Horse Lotteries:	The set H of functions $h : S \to O$.
Preferences:	A preference relation $\succeq$ on H.

Anscombe and Aumann (1963).

Subjective Expected Utility

SEU 1. $\succeq$ is complete and transitive.

SEU 2. Independence: If $h \succeq k$ then for all g and $0 \leq \alpha \leq 1$,
$\alpha g + (1-\alpha)h \succeq \alpha g + (1-\alpha)k$.

SEU 3. Archimedean axiom.

SEU 4. State independence.

There is a payoff function $u : O \to \mathbf{R}$ and a unique probability distribution p on S which define a functional V on H,

$$V(h) = \sum_s p(s) \sum_o u(o)h(s)(o)$$

such that $h \succeq k$ iff $V(h) \geq V(k)$.

Probabilistic Sophistication

Savage framework: S, O, acts $f : s \mapsto o$. P_0 set of finite-support probabilities on O

Definition: An individuali s said to be probabilistically sophisticated if there exists a probability measure p on S and a preference functional $V(x_1, p_1, \ldots, x_m, p_m)$ on P_0 satisfying mixture continuity and monotonicity with respect to stochastic dominance, such that preferences on acts are represented by the functional $f \mapsto V\big(x_1, p(f^{-1}(x_1)), \ldots, x_n, p(f^{-1}(x_n)\big)$ where $\{x_1, \ldots, x_n\}$ is the range of f.

Probabilistic Sophistication

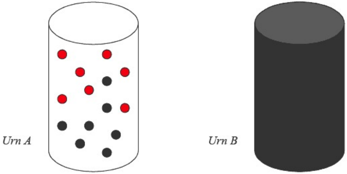
Savage framework: S, O, acts $f : s \mapsto o$. P_0 set of finite-support probabilities on O

Definition: An individuali s said to be probabilistically sophisticated if there exists a probability measure p on S and a preference functional $V(x_1, p_1, \ldots, x_n, p_n)$ on P_0 satisfying mixture continuity and monotonicity with respect to stochastic dominance, such that preferences on acts are represented by the functional $f \mapsto V(x_1, p(f^{-1}(x_1)), \ldots, x_n, p(f^{-1}(x_n))$ where $\{x_1, \ldots, x_n\}$ is the range of f.

- The separation of tastes and beliefs.
- Mixture-closed,
- Stochastic dominance: p stochastically dominates q iff for every increasing function f, $\int f\, dp \geq \int f\, dq$.

Why Non-Additive Probabilities?

- Ellsberg's Urns
- Schmeidler's Coins



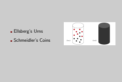


Image: http://scholar.lib.vt.edu/ejournals/SPT/v8n2/grinbaum.html.

Choquet Expected Utility

Definition: Acts f and g are comonotonic if $f(s) \succ f(t)$ implies $g(s) \succ g(t)$ for all $s, t \in S$.

CEU 1. $\succeq$ is complete and transitive.

CEU 2. Comonotonic independence.

CEU 3. Archimedean axiom.

CEU 4. State independence.

There is a payoff function $u : O \to \mathbf{R}$ and a unique capacity ϕ on S which define a functional V on H,

$$V(h) = \int_S \sum_o u(o)h(s)(o)d\phi(s)$$

such that $h \succeq k$ iff $V(h) \geq V(k)$.

MMEU

MMEU 1. $\succeq$ is complete and transitive.

MMEU 2. Certainty independence.

MMEU 3. Archimedean axiom.

MMEU 4. State independence.

MMEU 5. $f \sim g$ implies that for all $0 < \alpha < 1$, $\alpha f + (1 - \alpha) g \succeq f$.

There is a payoff function $u : O \to \mathbf{R}$ and a set P of probability distributions on S which define a functional V on H,

$$V(h) = \inf_{p \in P} \sum_s p(s) \sum_o u(o) h(s)(o)$$

such that $h \succeq k$ iff $V(h) \geq V(k)$.

MMEU

MMEU 1. $\succeq$ is complete and transitive.
MMEU 2. Certainty independence.
MMEU 3. Archimedean axiom.
MMEU 4. State independence.
MMEU 5. $f \sim g$ implies that for all $0 < \alpha < 1$, $\alpha f + (1-\alpha)g \succeq f$.

There is a payoff function $u : O \to \mathbf{R}$ and a set P of probability distributions on S which define a functional V on H,

$$V(h) = \inf_{p \in P} \sum_s p(s) \sum_o u(o)h(s)(o)$$

such that $h \succeq k$ iff $V(h) \geq V(k)$.

- Connection between MMEU, CEU. The core of a capacity is the set of all probability measures that dominate it. If ϕ is convex, $\phi(A \cup B) + \phi(A \cap B) \leq \phi(A) + \phi(B)$, then the Choquet integral of any real-valued function f with respect to ϕ is the minimum of the integrals with respect to probability distributions in the core.
- Notice that with equality, this is inclusion/exclusion, and is always satisfied by any probability measure.

The Problem with Conditioning

Savage Conditioning: $f \succeq_A g$ iff for some h, $f|_A h \succeq g|_A h$.

In SEU, any h gives the same answer, and $\succeq_A$ is represented by u and $p(\cdot|A)$.

The Problem with Conditioning

Savage Conditioning: $f \succeq_A g$ iff for some h, $f|_A h \succeq g|_A h$.

$S = \{x, y, z\}$. $\phi(x) = 1/4$, $\phi(y) = \phi(z) = 1/8$, $\phi(x, z) = 1/2$, $\phi(y, z) = 3/4$, $\phi(x, y) = 1/2$. Let $A = \{x, y\}$

$$f(x) = 1 \text{ on } x, \text{ else } 0, \qquad g(x) = 1 \text{ on } y, \text{ else } 0.$$

The Problem with Conditioning

Savage Conditioning: $f \succeq_A g$ iff for some h, $f|_A h \succeq g|_A h$.

$S = \{x, y, z\}$. $\phi(x) = 1/4$, $\phi(y) = \phi(z) = 1/8$, $\phi(x, z) = 1/2$, $\phi(y, z) = 3/4$, $\phi(x, y) = 1/2$. Let $A = \{x, y\}$

$f(x) = 1$ on x, else 0, $\qquad g(x) = 1$ on y, else 0.

$$\int f \, d\phi = 0 + 1\phi(x) = 1/4 \qquad \int g \, d\phi = 0 + 1\phi(y) = 1/8$$

The Problem with Conditioning

Savage Conditioning: $f \succeq_A g$ iff for some h, $f|_A h \succeq g|_A h$.

$S = \{x, y, z\}$. $\phi(x) = 1/4$, $\phi(y) = \phi(z) = 1/8$, $\phi(x, z) = 1/2$, $\phi(y, z) = 3/4$, $\phi(x, y) = 1/2$. Let $A = \{x, y\}$

$$f(x) = 1 \text{ on } x, \text{ else } 0, \qquad g(x) = 1 \text{ on } y, \text{ else } 0.$$

$$\int f\, d\phi = 0 + 1\phi(x) = 1/4 \qquad \int g\, d\phi = 0 + 1\phi(y) = 1/8$$

$h = 0$

$$\int f|_A h\, d\phi = 0 + 1\phi(x) = 1/4 \qquad \int g|_A h\, d\phi = 0 + 1\phi(y) = 1/8$$

The Problem with Conditioning

Savage Conditioning: $f \succeq_A g$ iff for some h, $f|_A h \succeq g|_A h$.

$S = \{x, y, z\}$. $\phi(x) = 1/4$, $\phi(y) = \phi(z) = 1/8$, $\phi(x, z) = 1/2$, $\phi(y, z) = 3/4$, $\phi(x, y) = 1/2$. Let $A = \{x, y\}$

$$f(x) = 1 \text{ on } x, \text{ else } 0, \qquad g(x) = 1 \text{ on } y, \text{ else } 0.$$

$$\int f\, d\phi = 0 + 1\phi(x) = 1/4 \qquad \int g\, d\phi = 0 + 1\phi(y) = 1/8$$

$h = 0$

$$\int f|_A h\, d\phi = 0 + 1\phi(x) = 1/4 \qquad \int g|_A h\, d\phi = 0 + 1\phi(y) = 1/8$$

$h = 1$

$$\int f|_A h\, d\phi = 0+1\phi(x, z) = 1/2 \qquad \int g|_A h\, d\phi = 0+1\phi(y, z) = 3/4$$

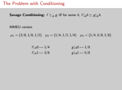
The Problem with Conditioning

Savage Conditioning: $f \succeq_A g$ iff for some h, $f|_A h \succeq g|_A h$.

MMEU version

$$\mu_1 = (3/8, 1/8, 1/2) \quad \mu_2 = (1/4, 1/2, 1/4) \quad \mu_3 = (1/4, 5/8, 1/8)$$

$$f|_A 0 \mapsto 1/4 \qquad g|_A 0 \mapsto 1/8$$
$$f|_A 1 \mapsto 3/8 \qquad g|_A 1 \mapsto 5/8$$



- Updating is the problem of defining conditional preference.
- Define $f_A h$.
- Same issue arises with MMEU. This is what I will talk about most.

The Conditioning Tradeoff

Dynamic Consistency: If $f \succ_A g$ in all ex post situations A, then $f \succeq g$.

Consequentialism: Conditional preferences given $A \succeq_A$ only depend on what happens in A.

Fact of Life: If the CEU updating rule satisfies dynamic consistency and conequentialism, then the capacity is a probability (and updating is Bayes).

f-Bayesian Updating

Definition: An updating rule is a map $(\succeq, A) \to \succeq_A$ that assigns to an unconditional preference relation and an event A the preference relation given A.

Definition: For act f, the f-Bayesian updating rule is $g \succeq_A^f h$ iff $g|_A f \succeq h_A f$.

f-Bayesian Updating

Definition: An updating rule is a map $(\succeq, A) \to \succeq_A$ that assigns to an unconditional preference relation and an event A the preference relation given A.

Definition: For act f, the f-Bayesian updating rule is $g \succeq_A^f h$ iff $g|_A f \succeq h_A f$.

The optimistic rule: $f \equiv o_*$. A is good news.

$$\phi_A(B) = \phi(B \cap A)/\phi(A).$$

The Bayesian rule.

Select and update all priors assigning maximal probability to A.

f-Bayesian Updating

Definition: An updating rule is a map $(\succeq, A) \to \succeq_A$ that assigns to an unconditional preference relation and an event A the preference relation given A.

Definition: For act f, the f-Bayesian updating rule is $g \succeq_A^f h$ iff $g|_A f \succeq h_A f$.

The pessimistic rule: $f \equiv o^*$. A is bad news.

$$\phi_A(B) = \left[\phi\big((B \cap A) \cup A^c\big) - \phi(A^c)\right] / \big(1 - \phi(A^c)\big).$$

The Dempster-Shafer rule for belief functions.

f-Bayesian Updating

Definition: An updating rule is a map $(\succeq, A) \to \succeq_A$ that assigns to an unconditional preference relation and an event A the preference relation given A.

Definition: For act f, the f-Bayesian updating rule is $g \succeq_A^f h$ iff $g|_A f \succeq h_A f$.

The pessimistic rule: $f \equiv o^*$. A is bad news.

$$\phi_A(B) = \left[\phi((B \cap A) \cup A^c) - \phi(A^c)\right] / (1 - \phi(A^c)).$$

The Dempster-Shafer rule for belief functions.

- Affine u.
- $|S| \geq 4$.
- Fagin-Halpern is full updating.

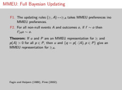
MMEU: Full Bayesian Updating

F1. The updating rules $(\succeq, A) \mapsto \succeq_A$ takes MMEU preferences ino MMEU preferences.

F2. For all non-null events A and outcomes o, if $f \sim o$ then $f|_A o \sim o$.

Theorem: If u and P are an MMEU representation for $\succeq$ and $p(A) > 0$ for all $p \in P$, then u and $\{q = p(\cdot|A), p \in P\}$ give an MMEU representation for $\succeq_A$.

Fagin and Halpern (1989), Pires (2002).



- MMEU is inherently pessimist, selecting priors which put the most weight on the worst outcomes. F2 guarantees that no matter the weights on E vs. E^c, the relative weights on states within E have to cohere with the unconditional weights.

Easy Implications: Portfolios

0 net positions in portfolios over a range of prices.

An asset pays off 1 in state H and 3 instate T. Trader beliefs are $\phi(H) = 0.3$ and $\phi(T) = 0.4$. The expected payoff of a unit long position at price p is

$$v_b = (1 - p) + 0.4(2) = 1.8 - p.$$

The value of a unit short position is

$$v_s = p - 3 + 0.3(2) = p - 2.4.$$

For $1.8 < p < 2.4$, the 0 position is better than both.

Home Bias Paradox
Equity Premium Puzzle

Easy Implications: Portfolios

0 net positions in portfolios over a range of prices.

An asset pays off 1 in state H and 3 instate T. Trader beliefs are $\phi(H) = 0.3$ and $\phi(T) = 0.4$. The expected payoff of a unit long position at price p is

$$v_b = (1 - p) + 0.4(2) = 1.8 - p.$$

The value of a unit short position is

$$v_s = p - 3 + 0.3(2) = p - 2.4.$$

For $1.8 < p < 2.4$, the 0 position is better than both.

Home Bias Paradox
Equity Premium Puzzle

- WIth SEU preferences, the 0 zone is a point, and there is indifference between buying and selling.

FINIS!

Definition: Non-Additive Probability

Definition: A non-additive probability ν on S is a function mapping subsets of S to $[0,1]$ such that

N.1. $\nu(\emptyset) = 0$,

N.2. $\nu(S) = 1$,

N.3. If $A \subset B$, then $\nu(A) \leq \nu(B)$.

For example, $S = \{s_1, s_2\}$.

$$\begin{aligned} \nu_\alpha(\emptyset) &= 0 \\ \nu_\alpha(s_1) = \nu_\alpha(s_2) &= \alpha \\ \nu_\alpha(S) &= 1 \end{aligned}$$

Definition: Non-Additive Probability

Definition: A non-additive probability ν on S is a function mapping subsets of S to $[0, 1]$ such that

N.1. $\nu(\emptyset) = 0$,

N.2. $\nu(S) = 1$,

N.3. If $A \subset B$, then $\nu(A) \leq \nu(B)$.

Integration

Suppose the values of f are $x_1 < \cdots < x_n$. Then

$$E_\nu f = x_1 + (x_2 - x_1)\nu(f > x_1) + \cdots (x_n - x_{n-1})\nu(f > x_{n-1}).$$

Dempster-Shafer Belief Functions

Definition: A belief function β on S is a function mapping subsets of S to $[0,1]$ such that

B.1. $\beta(\emptyset) = 0$,

B.2. $\beta(S) = 1$,

B.3. $\beta(\cup_{i=1}^{n} A_i) \geq \sum_{i=1}^{n} \sum_{I \subseteq \{1,\ldots,n\}:|I|=i} (-1)^{i+1} \beta(\cap_{j \in I} A_j)$.

Dempster-Shafer Belief Functions

Definition: A belief function β on S is a function mapping subsets of S to $[0, 1]$ such that

B.1. $\beta(\emptyset) = 0$,

B.2. $\beta(S) = 1$,

B.3. $\beta(\cup_{i=1}^{n} A_i) \geq \sum_{i=1}^{n} \sum_{I \subseteq \{1,\ldots,n\}:|I|=i} (-1)^{i+1} \beta(\cap_{j \in I} A_j)$.

B.3 is like inclusion-exclusion:

- $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$,
- $\Pr(A \cup B \cup C) = \Pr(A) + \Pr(B) + \Pr(C) - \Pr(A \cap B) - \Pr(A \cap C) - \Pr(B \cap C) + \Pr(A \cap B \cap C)$.

Dempster-Shafer Belief Functions

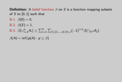
Definition: A belief function β on S is a function mapping subsets of S to $[0,1]$ such that

B.1. $\beta(\emptyset) = 0$,

B.2. $\beta(S) = 1$,

B.3. $\beta(\cup_{i=1}^{n} A_i) \geq \sum_{i=1}^{n} \sum_{I \subseteq \{1,\ldots,n\}:|I|=i} (-1)^{i+1} \beta(\cap_{j \in I} A_j)$.

$\beta(A) = \inf\{p(A) : p \geq \beta\}$



Definition: A belief function β on S is a function mapping subsets of S to $[0,1]$ such that

B.1. $\beta(\emptyset) = 0$,

B.2. $\beta(S) = 1$,

B.3. $\beta(\cup_{i=1}^{n} A_i) \geq \sum_{i=1}^{n} \sum_{I \subseteq \{1,\ldots,n\}:|I|=i} (-1)^{i+1} \beta(\cap_{j \in I} A_j)$.

$\beta(A) = \inf\{p(A) : p \geq \beta\}$

- Belief functions are tight capacities.
- $m(A)$ is the weight of evidence for A not assigned to any of its subsets.
- Theorem due to Shafer.

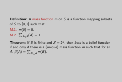
Mass Functions & Belief Functions

Definition: A mass function m on S is a function mapping subsets of S to $[0,1]$ such that

M.1. $m(\emptyset) = 0$,

M.2. $\sum_{A \in \mathcal{S}} (A) = 1$.

Theorem: If S is finite and $\mathcal{S} = 2^S$, then *beta* is a belief function if and only if there is a (unique) mass function m such that for all A, $\beta(A) = \sum_{B \subset A} m(B)$.



Definition: A mass function m on S is a function mapping subsets of S to $[0,1]$ such that

M.1. $m(\emptyset) = 0$,

M.2. $\sum_{A \subseteq S}(A) = 1$.

Theorem: If S is finite and $\mathcal{S} = 2^S$, then *beta* is a belief function if and only if there is a (unique) mass function m such that for all A, $\beta(A) = \sum_{B \subseteq A} m(B)$.

- Belief functions are tight capacities.
- $m(A)$ is the weight of evidence for A not assigned to any of its subsets.
- Theorem due to Shafer.

Definition: Lexicographic Probability

Definition: A lexicographic probability on S is a vector of probabilities $\mu = (\mu_1, \ldots, \mu_n)$ on S such that $A \succeq B$ iff the vector $\mu(A)$ lexicographically dominates $\mu(B)$.

LSEU

$$U(f) = \left(\mathrm{E}_{\mu_i} \left\{ \sum_o u(o) f(s)(o) \right\} \right)_{i=1}^{n}.$$

$f \succeq g$ if and only if $U(f) \geq U(g)$.

Blume Brandenburger and Dekel (1993a,b)

References

Blume, L., A. Brandenburger and E. Dekel (1993). Lexicographic probabilities and choice under uncertainty. *Econometrica*.

Cerreia-Vioglio,S., et. al., (2008). Uncertainty averse preferences. Unpublished.

Fagin, R. and Halpern, J. (1991). A new approach to updating belief, in Uncertainty in Artificial Intelligence 6.

Gelman, A. (2006). The boxer, the wrestler, and the coin flip: A paradox of robust Bayesian inference and belief functions. *The American Statistician*.

Gilboa, I. and D. Schmeidler (1989). Maxmin expected utility with a non-unique prior. *Journal of Mathematical Economics*.

Gilboa, I. and D. Schmeidler (1993). Updating ambiguous beliefs. *Journal of Economic Theory* 59.

Kast, R., A. Lapied and P. Toquebeuf (2008). Updating Choquet integrals, consequentialism and dynamic consistency. Unpublished.

Machina, M. and D. Schmeidler (1992). A more robust definition of subjective probability *Econometrica*.

Marinacci, M. and L. Montrucchio (2004). Introduction to the mathematics of ambiguity. In *Uncertainty in Economic Theory: A collection of essays in honor of David Schmeidlers 65th birthday.*

Pires, C (2002). A rule for updating ambiguous beliefs. *Theory and Decision*

Schmeidler, D. (1989). Subjective probability and expected utility without additivity, *Econometrica*.