Models and Decisions

Massimo Marinacci

AXA-Bocconi Chair in Risk
Department of Decision Sciences and IGIER
Università Bocconi

IMSI, University of Chicago
July 2021

Road map

- **Decision problems**
  - toolbox
  - the Savage and Anscombe-Aumann setups
  - classical subjective expected utility
- Model uncertainty: ambiguity / robustness models
- Issues
  - ambiguity / robustness makes optimal actions more prudent?
  - ambiguity / robustness favors diversification?
  - ambiguity / robustness affects valuation?
  - model ambiguity resolves in the long run through learning?
  - sources of uncertainty: a Pandora's box?
- Model misspecification

Probability of facts and of theories

- Decisions' consequences depend on external factors (contingencies)
- Probability of contingencies
- Probabilistic theories on contingencies (e.g., generative mechanisms, DGP)
- Thinking over such theories
- Two layers of uncertainty

Decision problems: the toolbox, I

A decision problem consists of

- a space A of actions
- a space C of material (e.g., monetary) consequences
- a space S of environment states
- a consequence function $\rho : A \times S \longrightarrow C$ that details the consequence
$$c = \rho(a, s)$$
of action a when state s obtains

Decision problems: the toolbox, I

- States are jointly exhaustive and mutually exclusive
- We thus abstract from state misspecification issues (e.g., unforeseen contingencies)

Example (i): natural hazards

Public officials have to decide whether or not to evacuate an area because of a possible earthquake

- A two actions a_0 (no evacuation) and a_1 (evacuation)
- C monetary consequences (damages to infrastructures and human casualties; Mercalli-type scale)
- S possible peak ground accelerations (Richter-type scale)
- $c = \rho\,(a, s)$ the monetary consequence of action a when state s obtains

Example (ii): monetary policy example

- ECB or the FED have to decide some target level of inflation to control the economy unemployment and inflation

- Unemployment u and inflation π outcomes are connected to shocks $\varepsilon = (\varepsilon_u, \varepsilon_\pi)$ and the policy a according to

$$
\begin{aligned}
u &= \theta_0 + \theta_{1\pi}\pi + \theta_{1a}a + \varepsilon_u \\
\pi &= a + \varepsilon_\pi
\end{aligned}
$$

- $\theta = (\theta_0, \theta_{1\pi}, \theta_{1a})$ are three structural coefficients

  (i) $\theta_{1\pi}$ and θ_{1a} are slope responses of unemployment to actual and planned inflation (e.g., Lucas-Sargent $\theta_{1a} = -\theta_{1\pi}$; Samuelson-Solow $\theta_{1a} = 0$)

  (ii) θ_0 is the rate of unemployment that would (systematically) prevail without policy interventions

Example (ii): monetary policy

Here:

A the target levels of inflation

C the pairs $c = (u, \pi)$

S has random and structural components

$$s = (\varepsilon, \theta)$$

The reduced form is

$$\begin{aligned} u &= \theta_0 + (\theta_{1\pi} + \theta_{1a})\, a + \theta_{1\pi}\varepsilon + \varepsilon_u \\ \pi &= a + \varepsilon_\pi \end{aligned}$$

and so ρ has the form

$$\rho(a, w, \varepsilon, \theta) = \begin{bmatrix} \theta_0 \\ 0 \end{bmatrix} + a \begin{bmatrix} \theta_{1\pi} + \theta_{1a} \\ 1 \end{bmatrix} + \begin{bmatrix} 1 & \theta_{1\pi} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_u \\ \varepsilon_\pi \end{bmatrix}$$

Example (ii): monetary policy

- Random components: shocks (i.e., minor omitted explanatory variables which we are "unable and unwilling to specify") or measurement errors
- Cf. the works of Hurwicz, Koopmans and Marschak in the 1940s and 1950s

Example (iii): climate policy

- A policy maker has to decide some target greenhouse gas emissions level to control damages associated with global temperatures increases
- Different sources of uncertainty are relevant

Example (iii): climate policy

- *Scientific uncertainty*: how do emissions E translate in increases of temperatures T? Assume

$$T = \theta_T E + \varepsilon_T$$

where θ_T is a structural CCR (carbon-climate response) parameter and ε_T is a random component

- *Socioeconomic uncertainty*: how do increases of temperatures T translate in economic damages D? Assume a DICE quadratic

$$D = \theta_{1D} T + \theta_{2D} T^2 + \varepsilon_D$$

where θ_{1D} and θ_{2D} are structural parameters and ε_D is a random component

- We abstract from issues about the objective functions

Example (iii): climate policy

Here:

A emission policies

C the economic damages (in GDP terms)

S has random and structural components

$$s = (\varepsilon, \theta)$$

where

$$\varepsilon = (\varepsilon_T, \varepsilon_D)$$

are the random components affecting the climate and economic systems, and

$$\theta = (\theta_T, \theta_{1D}, \theta_{2D})$$

are their structural coefficients

Example (iii): climate policy

- Action a is an emission policy, with cost $c(a)$
- $d(a, \varepsilon, \theta)$ economic damage function
- $\rho(a, \varepsilon, \theta) = -d(a, \varepsilon, \theta) - c(a)$ is the overall consequence of policy a
- From

$$\begin{cases} T = \theta_T a + \varepsilon_T \\ D = \theta_{1D} T + \theta_{2D} T^2 + \varepsilon_D \end{cases}$$

it follows that

$$\begin{aligned} d(a, \varepsilon, \theta) &= -(\theta_{1D}\theta_T + 2\theta_{2D}\varepsilon_T)\, a - \theta_{2D}\theta_T^2 a^2 - \theta_{1D}\varepsilon_T \\ &\quad -\theta_{2D}\varepsilon_T^2 - \varepsilon_D \end{aligned}$$

- More on this topic from Michael Barnett and William Brock

Decision problems: the toolbox, II

- The quartet (A, S, C, ρ) is a *decision form under uncertainty*
- The decision maker (DM) has a preference $\succsim$ over actions
  - we write $a \succsim b$ if the DM (weakly) prefers action a to action b
- The quintet $(A, S, C, \rho, \succsim)$ is a *decision problem under uncertainty*
- DMs aim to select actions $\hat{a} \in A$ such that $\hat{a} \succsim a$ for all $a \in A$
- Static setting, we abstract from temporal/dynamic issues

Consequentialism and the Savage setup

- What matters about actions is not their label / name but the *consequences* that they determine when the different states obtain
- *Consequentialism*: two actions that are realization equivalent – i.e., that generate the same consequence in every state – are indifferent
- We abstract from ethical issues

Consequentialism and the Savage setup

- Formally,

$$\rho(a, s) = \rho(b, s) \quad \forall s \in S \Longrightarrow a \sim b$$

or, equivalently,

$$\rho_a = \rho_b \Longrightarrow a \sim b$$

- Here $\rho_a : S \to C$ is the section of ρ at a given by $\rho_a(s) = \rho(a, s)$

The Savage setup

- The section ρ_a is a Savage *act*
- We can define a preference $\succsim$ over Savage acts by:

$$\rho_a \succsim \rho_b \iff a \succsim b$$

- For convenience, we keep using the same symbol $\succsim$

The Savage setup

- Savage's acts are typically denoted by $f : S \to C$
- The collection of all acts is denoted by $\mathcal{F}$
- The quartet $(\mathcal{F}, S, C, \succsim)$ is a *Savage decision problem under uncertainty*
- Through acts f_c constant to $c \in C$, i.e.

$$f_c(s) = c \qquad \forall s \in S$$

  the preference $\succsim$ induces a preference over consequences:

$$c \succsim c' \iff f_c \succsim f_{c'}$$

- Savage's setup is theoretically convenient but, in applications, acts may have a contrived interpretation

Random consequences

- In some applications, we are not able to specify an exhaustive state space
- A possibility is to assume that actions deliver consequences that are stochastic and not deterministic
- The consequence of action a is then a (finitely supported) probability distribution

  $$\rho(a) \in \Delta_0(C)$$

  on consequences, called *lottery*
- We denote by p and q typical lotteries; for each lottery p, the quantity

  $$p(c) \in [0,1]$$

  is the probability that consequence c obtains

Random consequences

- We identify a consequence $c \in C$ with the trivial (Dirac) lottery δ_c that assigns probability 1 to c, i.e.,

$$\delta_c\left(c'\right) = \begin{cases} 1 & \text{if } c' = c \\ 0 & \text{else} \end{cases}$$

- Up to this identification, we can regard C as a subset of $\Delta_0\left(C\right)$

Random consequentialism and the Anscombe-Aumann setup

- Consider action with random consequences, i.e., lotteries
- *Random consequentialism*: two actions sharing the same random consequence in every state are indifferent
- Formally,

$$\rho(a, s) = \rho(b, s) \quad \forall s \in S \Longrightarrow a \sim b$$

or, equivalently,

$$\rho_a = \rho_b \Longrightarrow a \sim b$$

- Random consequentialism subsumes (outcome) consequentialism: recall the identifications of consequences and trivial lotteries
- The section ρ_a is a an Anscombe-Aumann (AA) *act*

The Anscombe-Aumann setup

- AA acts are defined by $f : S \longrightarrow \Delta_0(C)$
- The collection of all acts is denoted by $\mathcal{F}$
- The quartet $(\mathcal{F}, S, C, \succsim)$ is an *AA decision problem under uncertainty*
- As in the Savage's setup, through constant acts the preference $\succsim$ induces a preference over lotteries
- Through trivial lotteries, in turn this preference over lotteries induces a preference over non-random consequences:

$$c \succsim c' \Longleftrightarrow \delta_c \succsim \delta_{c'}$$

The Anscombe-Aumann setup

- The AA consequence space has a vector structure – often in place of $\Delta_0(C)$ one considers a convex subset of a vector space
- By mixing AA acts
$$\alpha f + (1-\alpha) g$$
with
$$(\alpha f + (1-\alpha) g)(s) = \alpha f(s) + (1-\alpha) g(s) \qquad \forall s \in S$$
the space $\mathcal{F}$ inherits this vector structure, a very convenient feature of the AA setup, widely used in the theoretical literature
- Yet, the interpretation of mixing (often via randomization) can be contrived

Probability models

- Because of their ex-ante structural information, DMs know that states are generated by a probability model m that belongs to a given subset M of $\Delta(S)$
- Each m describes a possible DGP, so it represents (model) risk
- DMs thus posit a model space M in addition to the state space S, a central tenet of classical statistics a la Neyman-Pearson-Wald
- When the model space is based on experts' advice, its nonsingleton nature may reflect different advice

Models: a toy example

Consider an urn with 90 Red, or Green, or Yellow balls

- DMs bet on the color of a ball drawn from the urn
- State space is $S = \{R, G, Y\}$
- Without any further information, $M = \Delta(\{R, G, Y\})$
- If DMs are told that 30 balls are red, then

$$M = \left\{ m \in \Delta(\{R, G, Y\}) : m(R) = \frac{1}{3} \right\}$$

Models and experts: probability of heart attack

Two DMs: John and Lisa are 70 years old

- smoke
- no blood pressure problem
- total cholesterol level 310 mg/dL
- HDL-C (good cholesterol) 45 mg/dL
- systolic blood pressure 130

What's the probability of a heart attack in the next 10 years?

Models and experts: probability of heart attack

Based on their data and medical models, experts say

Experts	John's m	Lisa's m
Mayo Clinic	25%	11%
National Cholesterol Education Program	27%	21%
American Heart Association	25%	11%
Medical College of Wisconsin	53%	27%
University of Maryland Heart Center	50%	27%

Table from Gilboa and M. (2013)

Uncertainty: a taxonomy

In this setup, we can decompose uncertainty in three distinct layers:

- *Model risk*: uncertainty within a model m
- *Model ambiguity*: uncertainty across models in M
- *Model misspecification*: uncertainty about models (the correct model does not belong to the posited set M)

Models: a consistency condition

- Cerreia-Vioglio et al. (2013) take the "structural" information M as a primitive and thus enrich the standard framework
- DMs know that the correct model m that generates observations belongs to the posited collection M
- In terms of preferences: betting behavior must be *consistent* with datum M, i.e.,

$$m(F) \geq m(E) \quad \forall m \in M \Longrightarrow \text{"bet on } F\text{"} \succsim \text{"bet on } E\text{"}$$

- The sextet $(A, S, C, M, \rho, \succsim)$ forms a *classical decision problem under uncertainty*
- Here we abstract from model misspecification issues (to be dealt with later)

Risk: EU

- Suppose that the DMs know the correct model m, so M is a singleton
- A preference $\succsim$ that satisfies Savage's axioms and the consistency condition is represented by the *expected utility* criterion
$$V(a) = \sum_s u(\rho(a,s))\, m(s)$$
That is, actions a and b are ranked as follows:
$$a \succsim b \Longleftrightarrow V(a) \geq V(b)$$
- u is a von Neumann-Morgenstern utility function:
$$c \succsim c' \Longleftrightarrow u(c) \geq u(c')$$
- It captures *risk attitudes*

Model ambiguity: classical SEU

A preference $\succsim$ that satisfies Savage's axioms and the consistency condition is represented by the *classical subjective expected utility* (*SEU*) criterion

$$V(a) = \sum_m \left(\sum_s u(\rho(a, s))\, m(s) \right) \mu(m)$$

That is, actions a and b are ranked as follows:

$$a \succsim b \iff V(a) \geq V(b)$$

Here

- u is again a von Neumann-Morgenstern utility function
- μ is a *subjective prior probability* that quantifies the uncertainty about models; its support is included in M
- If M is based on the advice of different experts, the prior may reflect the *different confidence* that DMs have in each of them

Model ambiguity: classical SEU

- The "classical" adjective reminds of the classical statistics tenet on which this criterion relies
- If we set

$$\mathrm{R}(a, m) = \sum_s u(\rho(a, s))\, m(s)$$

we can write the classical SEU criterion as

$$V(a) = \sum_m \mathrm{R}(a, m)\, \mu(m)$$

- In words, this criterion considers the expected utility $\mathrm{R}(a, m)$ of each possible model m, and averages them out according to the prior μ

Model ambiguity: classical SEU

- Each prior μ induces a *predictive probability* $\bar{\mu} \in \Delta(S)$ through reduction

$$\bar{\mu}(E) = \sum_m m(E)\,\mu(m)$$

In turn, the predictive probability enables to rewrite the classical SEU criterion as

$$V(a) = \mathrm{R}(a, \bar{\mu}) = \sum_s u(\rho(a, s))\,\bar{\mu}(s)$$

- This reduced form of V is the original Savage subjective EU representation

Classical SEU: some special cases

- If the support of μ is a singleton $\{m\}$, DMs subjectively (and so possibly wrongly) believe that m is the correct model. The criterion thus reduces to a Savage EU criterion $\mathrm{R}(a, m)$
- If M is a singleton $\{m\}$, DMs know that m is the correct model (a rational expectations tenet)
  - (i) There is only model risk (quantified by m)
  - (ii) The criterion again reduces to the EU representation $\mathrm{R}(a, m)$, but now interpreted as a *von Neumann-Morgenstern criterion*

Classical SEU: some special cases

- Singleton M have been pervasive in economics
- Since the 70s, economics has emphasized the study of agents' reactions to the "opponents" actions (from the Lucas critique in macroeconomics to the study of incentives in game theoretic settings)
- Rational expectations literature had to depart from the "particle" view of agents of the Keynesian macroeconomics of the 50s and 60s

Factorization

- In applications, states often have random and structural components

$$s = (\varepsilon, \theta)$$

- The shock has the form

$$\varepsilon = \sigma w$$

where w is a "white noise" with zero mean and unit variance

- The parameter $\sigma \in \Sigma$ specifies the standard deviation of the shock
- DMs know the shock distribution, up to the standard deviations σ

Factorization

- The positive scalar

$$m\left(\theta, \varepsilon\right)$$

gives the joint probability of parameters and shocks under model m

- We consider models factored as:

$$m = \delta_\theta \times q_\sigma$$

i.e.,

$$m(\varepsilon, \theta') = \begin{cases} q_\sigma\left(\varepsilon\right) & \text{if } \theta' = \theta \\ 0 & \text{else} \end{cases}$$

- Each model corresponds to
  1. a distribution q_σ of the random component ε
  2. a parameter θ (e.g., a model climate system/economy)

Factorization

- In the factorization $m = q \times \delta_\theta$, two kinds of model uncertainties emerge:
- *Theoretical model ambiguity* about the economic and physical theories that underpin the models: different θ correspond to different theories
- *Stochastic model ambiguity* about the statistical performance of such theories, due to shocks and to measurement errors: different q_σ correspond to different performances

Factorization

- We write the consequence function as $\rho_\theta\left(a, \varepsilon\right)$ to emphasize the structural component θ over the random one ε
- We index factored models as

$$m_{\theta,\sigma} = q_\sigma \times \delta_\theta$$

- An hypothesis on states is summarized by a pair $\left(\theta, \sigma\right) \in \mathcal{H} \subseteq \Theta \times \Sigma$
- The set of models that the DM posits is

$$M = \left\{m_{\theta,\sigma} : \left(\theta, \sigma\right) \in \mathcal{H}\right\}$$

Factorization

- Model risk is within each q_σ
- Model ambiguity is over the structural coefficient θ and the standard deviation σ
- To address it, the DM has a prior probability $\mu(\theta, \sigma)$ that quantifies DM's degree of belief that θ is the true parameter

Classical SEU under factorization

- We have
$$\mathrm{R}\left(a, p\right) = \sum_{\theta,\varepsilon} u\left(\rho_\theta(a,\varepsilon)\right) p(\theta,\varepsilon)$$
for each $p \in \Delta$
- In particular, for a factored model indexed by a pair $(\theta, \sigma) \in \Theta \times \Sigma$ we have
$$\begin{aligned}
\mathrm{R}\left(a, \theta, \sigma\right) &= \sum_{\theta',\varepsilon} u\left(\rho_{\theta'}(a,\varepsilon)\right) m_{\theta,\sigma}(\theta',\varepsilon) \\
&= \sum_{\theta',\varepsilon} u\left(\rho_{\theta'}(a,\varepsilon)\right) \left(q_\sigma \times \delta_\theta\right)(\theta',\varepsilon) \\
&= \sum_{\varepsilon} u\left(\rho_\theta(a,\varepsilon)\right) q_\sigma(\varepsilon)
\end{aligned}$$

Classical SEU under factorization

- The classical SEU criterion becomes

$$V(a) = \sum_{\theta,\sigma} \left(\sum_{\varepsilon} u\left(\rho_\theta(a,\varepsilon)\right) q_\sigma(\varepsilon) \right) \mu(\theta,\sigma)$$

or, equivalently,

$$V(a) = \sum_{\theta,\sigma} \mathrm{R}(a,\theta,\sigma)\, \mu(\theta,\sigma)$$

Factorized classical SEU: monetary policy example

- Back to the monetary example

$$u = \theta_0 + \theta_{1\pi}\pi + \theta_{1a}a + \varepsilon_u$$
$$\pi = a + \varepsilon_\pi$$

- The shock $\varepsilon = (\varepsilon_u, \varepsilon_\pi)$ has the form

$$\varepsilon_u = \sigma_u w \quad \text{and} \quad \varepsilon_\pi = \sigma_\pi w'$$

where w and w' are uncorrelated "white noises" with zero mean and unit variance

- Distribution q_σ of shock ε is known up to the vector

$$\sigma = (\sigma_u, \sigma_\pi)$$

of standard deviations

Factorized classical SEU: monetary policy example

- Model economy θ is unknown
- So, belief μ is on (θ, σ)
- The *monetary policy problem* is then

$$\max_{a \in A} V(a) = \max_{a \in A} \sum_{\theta,\sigma} \left(\sum_{\varepsilon} u\left(\rho_{\theta}(a, \varepsilon)\right) q_{\sigma}(\varepsilon) \right) \mu(\theta, \sigma)$$

Road map

- Decision problems
  - toolbox
  - the Savage and Anscombe-Aumann setups
  - classical subjective expected utility
- **Model uncertainty: ambiguity / robustness models**
- Issues
  - ambiguity / robustness makes optimal actions more prudent?
  - ambiguity / robustness favors diversification?
  - ambiguity / robustness affects valuation?
  - model ambiguity resolves in the long run through learning?
  - sources of uncertainty: a Pandora's box?
- Model misspecification

Ambiguity / Robustness: the problem

- Model risk and ambiguity need to be treated differently
- The standard expected utility model does not
- Since the 1990s, a strand of economic literature has been studying *ambiguity* / *Knightian uncertainty* / *robustness* / *deep uncertainty*
- Normative focus (no behavioral biases or "mistakes")
- We consider two approaches
  - non-Bayesian (Gilboa and Schmeidler 1989)
  - Bayesian (Klibanoff, M. and Mukerji 2005)
- Both approaches broaden the scope of traditional EU analysis

Ambiguity / Robustness: the problem

- Intuition: betting on coins is greatly affected by whether or not coins are well tested
- Models correspond to possible biases of the coin
- By symmetry (uniform reduction), heads and tails are judged to be equally likely when betting on an untested coin, never flipped before
- The same probabilistic judgement holds for a well tested coin, flipped a number of times with an approximately equal proportion of heads to tails
- The evidence behind such judgements, and so the confidence in them, is dramatically different: ceteris paribus, DMs may well prefer to bet on tested (model risk) rather than on untested coins (model risk & ambiguity)
- Experimental evidence: Ellsberg paradox

Ambiguity / Robustness: relevance

- A more robust rational behavior toward uncertainty emerges
- A more accurate / realistic account of how uncertainty affects valuation (e.g., uncertainty premia in market prices)
- Better understanding of exchange mechanics
  - a dark side of uncertainty: no-trade or small-trade results because of cumulative effects of model risk and ambiguity; see the financial crisis
- Better calibration and quantitative exercises
  - applications in Finance, Macroeconomics, and Environmental Economics
- Better modelling of decision / policy making
  - applications in Risk Management; e.g., the otherwise elusive precautionary principle may fit within this framework

Ambiguity / Robustness: relevance

- Caveat: model risk and ambiguity can work in the same direction (magnification effects), as well as in different directions
- Magnification effects: large "uncertainty prices" with reasonable degrees of risk aversion
- Combination of sophisticated formal reasoning and empirical relevance

Ambiguity / Robustness: a Bayesian approach

- A first distinction: DMs do not have attitudes toward uncertainty per se, but rather toward model risk and model ambiguity
- Such attitudes may differ: typically DMs are more averse to model ambiguity than to model risk
- Experimental evidence from Aydogan et al. (2018)

Bayesian approach: a tacit assumption

- Suppose consequences are monetary
- Recall that $\mathrm{R}(a, m) = \sum_s u(\rho(a, s))\, m(s)$
- Classical subjective EU representation can be written as

$$\begin{aligned} V(a) &= \sum_m \mathrm{R}(a, m)\, \mu(m) \\ &= \sum_m \left(u \circ u^{-1}\right)(\mathrm{R}(a, m))\, \mu(m) \\ &= \sum_m u(c(a, m))\, \mu(m) \end{aligned}$$

where $c(a, m)$ is the certainty equivalent

$$c(a, m) = u^{-1}(\mathrm{R}(a, m))$$

of action a under model m

Bayesian approach: a tacit assumption

- The profile
$$\{c(a,m) : m \in \operatorname{supp} \mu\}$$
is the scope of the model ambiguity that is relevant for the decision
- In particular, DMs use the decision criterion
$$V(a) = \sum_m u(c(a,m))\,\mu(m)$$
to address model ambiguity, while
$$\mathrm{R}(a,m) = \sum_s u(\rho(a,s))\,m(s)$$
is how DMs address the model risk that each m features
- Identical attitudes toward model risk and ambiguity, both described by the same function u

Bayesian approach: representation

- The smooth ambiguity model generalizes the representation by distinguishing such attitudes
- Actions are ranked according to the *smooth ambiguity* criterion

$$\begin{aligned} V(a) &= \sum_m \left(v \circ u^{-1}\right)(\mathrm{R}(a,m))\,\mu(m) \\ &= \sum_m v(c(a,m))\,\mu(m) \end{aligned}$$

- The function $v : C \to \mathbb{R}$ represents attitudes toward model ambiguity

Bayesian approach: representation

- A negative attitude toward model ambiguity is modelled by a concave v, interpreted as aversion to (mean preserving) spreads in certainty equivalents $c\,(a, m)$
- Ambiguity aversion amounts to a higher degree of aversion toward model ambiguity than toward model risk, i.e., a v more concave than u

Bayesian approach: representation

- Setting $\phi = v \circ u^{-1}$, the smooth ambiguity criterion can be written as
$$V(a) = \sum_m \phi\left(\mathrm{R}(a, m)\right) \mu(m)$$
- This formulation holds for any kind of consequence (not just monetary)
- Ambiguity aversion corresponds to the concavity of ϕ, a "portable" feature
- If $\phi(x) = -e^{-\lambda x}$, it is a Bayesian version of the multiplier preferences (Hansen and Sargent 2001, 2008)
- Sources of uncertainty now matter – no longer "uncertainty is reduced to risk"

Bayesian approach: extreme attitudes and maxmin

- Under extreme ambiguity aversion (e.g., as $\lambda \uparrow \infty$ when $\phi(x) = -e^{-\lambda x}$), the smooth ambiguity criterion in the limit reduces to the maxmin criterion

$$V(a) = \min_{m \in \operatorname{supp} \mu} \sum_s u(\rho(a,s))\, m(s)$$

- Pessimistic criterion: DMs maxminimize over all possible probability models in the support of μ
- The prior μ just selects which models in M are relevant
- It is, essentially, the maxmin criterion of Wald (1950)
- Gilboa and Schmeidler (1989) seminal maxmin decision model can take a Waldean interpretation

Bayesian approach: extreme attitudes and maxmin

- If $\mathrm{supp}\,\mu = M$, the prior is actually irrelevant and we get back to a *stricto sensu* Wald maxmin criterion

$$V(a) = \min_{m} \sum_{s \in S} u(\rho(a, s))\, m(s)$$

- When M consists of all possible models, it reduces to the statewise maxmin criterion

$$V(a) = \min_{s} u(\rho(a, s))$$

  A very pessimistic (paranoid?) criterion: probabilities, of any sort, do not play any role (Arrow-Hurwicz decision under ignorance)

- Precautionary principle

Bayesian approach: remarks

- Under maxmin behavior there might be no trade on assets (Dow and Werlang, 1992). More generally, a lower trade volume on assets may correspond to a higher ambiguity aversion (e.g., higher λ when $\phi(x) = -e^{-\lambda x}$)

- So, ambiguity reinforces the idea that uncertainty can be an impediment to trade

- The smooth ambiguity criterion admits a simple quadratic approximation that generalizes the classic mean-variance model (Maccheroni, M. and Ruffino, 2013)

Ambiguity / Robustness: a non Bayesian approach

- Need to relax the requirement that a single number quantifies beliefs: the multiple (prior) probabilities model
- DMs may not have enough information to quantify their beliefs through a single probability, but need a set of them
- Expected utility is computed with respect to each probability and DMs act according to the minimum among such expected utilities

Non Bayesian approach: representation

- Model ambiguity addressed through a set C of priors
- DMs use the *multiple priors* criterion

$$\begin{aligned} V(\mathbf{a}) &= \min_{\mu \in C} \sum_m \left(\sum_s u(\rho(a,s))\, m(s) \right) \mu(m) \\ &= \min_{\mu \in C} \sum_s u(\rho(a,s))\, \bar{\mu}(s) \end{aligned} \tag{1}$$

- DMs consider the least among all the EU determined by each prior in C
- The predictive form (1) is the original version axiomatized by Gilboa and Schmeidler (1989)

Non Bayesian approach: comments

- This criterion is less extreme than it may appear at a first glance
- The set C incorporates
  - the attitude toward ambiguity, a taste component
  - its perception, an information component
- A smaller set C may reflect both better information – i.e., a lower perception of ambiguity – and / or a less averse ambiguity attitude
- In sum, the size of C does not reflect just information, but taste as well

Non Bayesian approach: comments

- With singletons $C = \{\mu\}$ we return to the classical subjective EU criterion

- When C consists of all possible priors on M, we return to the Wald maxmin criterion

$$\min_{m} \sum_{s} u(\rho(a, s))\, m(s)$$

- No trade results (kinks)

Non Bayesian approach: comments

A more general *α-maxmin criterion* has been axiomatized by Ghirardato, Maccheroni and M. (2004):

$$\begin{aligned} V(a) &= \alpha \min_{\mu \in C} \sum_m \left(\sum_s u\left(\rho\left(a, s\right)\right) m\left(s\right) \right) \mu(m) \\ &\quad + (1-\alpha) \max_{\mu \in C} \sum_m \left(\sum_s u\left(\rho\left(a, s\right)\right) m\left(s\right) \right) \mu(m) \end{aligned}$$

Non Bayesian approach: variational model

- In the multiple priors model, a prior μ is either "in" or "out" of the set C
- Maccheroni, M. and Rustichini (2006): general *variational* criterion

$$V(\mathbf{a}) = \inf_{\mu \in \Delta(M)} \left(\sum_m \left(\sum_s u(\rho(a,s))\, m(s) \right) \mu(m) + c(\mu) \right)$$

where $c(\mu)$ is a convex function that weights each prior μ

- If c is the dichotomic function given by

$$\delta_C(\mu) = \begin{cases} 0 & \text{if } \mu \in C \\ +\infty & \text{else} \end{cases}$$

we get back to the multiple priors model with set of priors C

Non Bayesian approach: multiplier model

- If c is given by the relative entropy $R(\mu||\nu)$, where ν is a reference prior, we get the *multiplier* criterion

$$V(a) = \inf_{\mu \in \Delta(M)} \left(\sum_m \left(\sum_s u(\rho(a,s))\, m(s) \right) \mu(m) + \alpha R(\mu||\nu) \right)$$

  popularized by Hansen and Sargent in their studies on robustness in Macroeconomics

- Also the mean-variance criterion is variational, with c given by a Gini index

Road map

- Decision problems
  - toolbox
  - the Savage and Anscombe-Aumann setups
  - classical subjective expected utility
- Model uncertainty: ambiguity / robustness models
- **Issues (skipped)**
  - ambiguity / robustness makes optimal actions more prudent?
  - ambiguity / robustness favors diversification?
  - ambiguity / robustness affects valuation?
  - model ambiguity resolves in the long run through learning?
  - dynamics: recursive models
  - sources of uncertainty: a Pandora's box?
- Model misspecification

Sources of uncertainty

- We made a distinction between attitudes toward model risk and model ambiguity
- A more general issue: do attitudes toward different uncertainties differ?
- Source contingent outcomes: do DMs regard outcomes (even monetary) that depend on different sources as different economic objects?
- Ongoing research on this subtle topic

Interim epilogue

- In decision problems with data, it is important to distinguish model risk, ambiguity and misspecification
- Traditional EU reduces model ambiguity to model risk, so it ignores the distinction
- Experimental and empirical evidence suggest that the distinction is relevant and may affect valuation
- We presented two approaches, one Bayesian and one not
- For different applications, different approaches may be most appropriate
- Model misspecification can be studied within this framework, as we will see next

Road map

- Decision problems
  - toolbox
  - Savage setup
  - classical subjective expected utility
- Model uncertainty: ambiguity / robustness models
- Issues
  - ambiguity / robustness makes optimal actions more prudent?
  - ambiguity / robustness favors diversification?
  - ambiguity / robustness affects valuation?
  - model ambiguity resolves in the long run through learning?
  - sources of uncertainty: a Pandora's box?
- **Model misspecification**

Decision making under model uncertainty

- Decisions' consequences depend on external factors (contingencies)
- Probability of contingencies
- Probabilistic theories on contingencies (e.g., generative mechanisms, DGP)
- Thinking over such theories
- Environments with uncertainty through the guise of models (e.g., policy making)
- Decision making under model uncertainty
- Based on Cerreia-Vioglio et al. (2021)

Setup

Recall that a Savage decision problem consists of

- a space $\mathcal{F}$ of acts $f : S \to C$
- a space C of material (e.g., monetary) consequences
- a space S of environment states
- The quartet $(\mathcal{F}, S, C, \succsim)$ is a *Savage decision problem under uncertainty*
- If C is a convex subset of a vector space (say, consisting of lotteries), this quartet takes the *Anscombe-Aumann* form
- We abstract from state misspecification issues (e.g., unforeseen contingencies)

Structured models

- Δ is the set of probability measures on S
- Recall that DMs posit a set M of *models* $m \in \Delta$ on states, with a substantive motivation or scientific underpinnings
- Each m describes a possible *DGP*, so it represents model risk
- Here it becomes convenient to call *structured* the models in M to emphasize their substantive motivation

Structured models

- DMs thus posit a model space M in addition to the state space S
- When the model space is based on experts' advice, its nonsingleton nature may reflect different advice
- If needed, M is a convex and compact subset of Δ^{σ}

The uncertainty taxonomy

- The quintet $(\mathcal{F}, S, C, M, \succsim)$ forms a *classical decision problem under uncertainty*
- If DMs know that the correct model belongs to M, they confront *model ambiguity*
- If DMs know the correct model within M, they confront *risk*

The uncertainty taxonomy

Recall that, in this setup, we can decompose uncertainty in three distinct layers:

- *Model risk*: uncertainty within a model m
- *Model ambiguity*: uncertainty across models in M
- *Model misspecification*: uncertainty about models (the correct model does not belong to the posited set M)

Model misspecification: Relevance

- Do data reveal DGPs and so speak, by and large, for themselves?
- If so, model misspecification is a minor issue
- Is theoretical reasoning needed to interpret empirical phenomena?
- If so, model misspecification is a major issue

Model misspecification: Issues

- Need of a decision criterion that accounts for model misspecification concerns
- Currently, models with agents confronting model misspecification are unable to address agents' misspecification concerns (they even use expected utility preferences)

Model misspecification

- Suppose that DMs confront model misspecification
- At the time of decision, they are afraid that none of the posited structured models is correct

Model misspecification (Hansen and Sargent, 2020)

- The DM contemplates also *unstructured models* $p \in \Delta$ in ranking actions according, for example, to a conservative decision criterion

$$V(f) = \min_{p \in \Delta} \left\{ \int u(f)\, dp + \lambda \min_{m \in M} R(p||m) \right\}$$

- $\lambda > 0$ is an index of misspecification fear
- The relative entropy $R(\cdot||\cdot)$ is an index of statistical distance between models (structured or not)
- So, $\min_{m \in M} R(p||m)$ is an Hausdorff "distance" between p and M
- We have $\min_{m \in M} R(p||m) > 0$ iff $p \notin M$

A protective belt

- Unstructured models lack the substantive status of structured models, they are essentially statistical artifacts
- In this variational criterion, they act as a protective belt against model misspecification

Model ambiguity: back to Wald 1950

- The higher λ is, the lower the misspecification fear is

- If $\lambda = +\infty$, the criterion takes a maxmin form

$$V(f) = \min_{m \in M} \int u(f)\, dm$$

  and we are back to model ambiguity

- Without misspecification fear, the DM would maxminimize over structured models

- No prior beliefs (cf. general maxmin analysis of Gilboa and Schmeidler, 1989)

Multiplier criterion

- If M is a singleton $\{m\}$, so no model ambiguity, we have the multiplier criterion

$$V(f) = \min_{p \in \Delta} \left\{ \int u(f)\, dp + \lambda R(p||m) \right\}$$

- Under the protective belt interpretation, it is the criterion of an expected utility DM who fears model misspecification (about the unique posited model)

General form

- In general, a decision criterion under model misspecification is

$$V(f) = \min_{p \in \Delta} \left\{ \int u(f)\, dp + \min_{m \in M} c(p, m) \right\}$$

- Here $c : \Delta \times M \to [0, \infty]$ is a statistical distance (for the set M), with $c(p, m) = 0$ iff $m = p$

- E.g., the relative entropy $R(\cdot||\cdot)$ or, more generally, a Csiszar ϕ-divergence $D_\phi(\cdot||\cdot)$

- We have $\min_{m \in M} c(p||m) > 0$ iff $p \notin M$

Box and all that

- Structured models may be incorrect, yet useful as Box (1979) famously remarked

- Formally, betting behavior must be *consistent* with datum M, i.e.,

$$m(F) \geq m(E) \quad \forall m \in M \Longrightarrow \text{“bet on } F\text{”} \succsim \text{“bet on } E\text{”}$$

- Under bet-consistency, a DM may fear model misspecification yet regards structured models as good enough to choose to bet on events that they unanimously rank as more likely

Mild model misspecification

- A mild form of fear of model misspecification

- **PROP** The decision criterion

$$V(f) = \min_{p \in \Delta} \left\{ \int u(f)\, dp + \lambda \min_{m \in M} R(p || m) \right\}$$

  is bet-consistent

- The result continues to hold for any ϕ-divergence $D_{\phi}(p||m)$

Misspecification neutrality

- A preference $\succsim$ is *misspecification neutral* if

$$\int u(f)\, dm \geq \int u(g)\, dm \quad \forall m \in M \Longrightarrow f \succsim g$$

  for all acts f and g

- In this case, for decision-theoretic purposes fear of misspecification plays no role

- We are back to aversion to model ambiguity

Misspecification neutrality

- **PROP** A preference $\succsim$ represented by the decision criterion

$$V(f) = \min_{p \in \Delta} \left\{ \int u(f)\, dp + \min_{m \in M} c(p, m) \right\}$$

  is misspecification neutral iff it is represented by the maxmin criterion

$$V(f) = \min_{m \in M} \int u(f)\, dm$$

- This confirms behaviorally that the maxmin criterion corresponds to aversion to model ambiguity, with no fear of misspecification.

A tale of two preferences

- This criterion can be axiomatized within a two-preference setup a la Gilboa et al. (2010), in an Anscombe-Aumann setting
- A dominance relation $\succsim^*$ represents the DM "genuine" preference on acts, so it is typically incomplete
- A behavioral preference $\succsim$ governs choice, so it is complete (burden of choice)

To be continued

- Bayesian analysis (unforeseen contingencies one level up)
- Dynamic analysis (more on this from Eran Hanany and Lars Peter Hansen)
- Applications

Slides' references

- I. Aydogan, L. Berger, V. Bosetti and N. Liu, Three layers of uncertainty and the role of model misspecification, mimeo, 2018
- G. E. P. Box, Science and statistics, *J. American Statistical Association*, 71, 791-799, 1976
- S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci and L. Montrucchio, Classical subjective expected utility, *PNAS*, 110, 6754-6759, 2013
- S. Cerreia-Vioglio, L. P. Hansen, F. Maccheroni and M., Making decisions under model misspecification, mimeo, 2021
- J. Dow and S. R. C. Werlang, Uncertainty aversion, risk aversion, and the optimal choice of portfolio, *Econometrica*, 60, 197-204, 1992

Slides' references

- P. Ghirardato, F. Maccheroni and M. Marinacci, Differentiating ambiguity and ambiguity attitude, *J. Economic Theory*, 118, 133-173, 2004
- I. Gilboa, F. Maccheroni, M. Marinacci and D. Schmeidler, Objective and subjective rationality in a multiple prior model, *Econometrica*, 78, 755-770, 2010
- I. Gilboa and D. Schmeidler, Maxmin expected utility with non-unique prior, *J. Mathematical Economics*, 18, 141-153, 1989
- L. P. Hansen and T. J. Sargent, Robust control and model uncertainty, *American Economic Review*, 91, 60-66, 2001
- L. P. Hansen and T. J. Sargent, *Robustness*, Princeton University Press, Princeton, 2008

Slides' references

- L. P. Hansen and T. J. Sargent, Structured ambiguity and model misspecification, 2020 (*J. Economic Theory*, forthcoming)
- P. Klibanoff, M. Marinacci and S. Mukerji, A smooth model of decision making under ambiguity, *Econometrica*, 73, 1849-1892, 2005
- F. Maccheroni, M. Marinacci and A. Rustichini, Ambiguity aversion, robustness, and the variational representation of preferences, *Econometrica*, 74, 1447-1498, 2006
- F. Maccheroni, M. Marinacci and D. Ruffino, Alpha as ambiguity: Robust mean-variance portfolio analysis, *Econometrica*, 81, 1075-1113, 2013.
- A. Wald, *Statistical decision functions*, Wiley, New York, 1950

Some readings

- Evergreen works from the founding fathers:
  1. F. Ramsey, *Truth and Probability*, 1926
  2. L. Savage, *Foundations of statistics*, 1954 (now a Dover book)
  3. B. de Finetti, *Teoria della probabilità*, 1970 (trans. 1974, Wiley)
- Classical presentations of the classical theory:
  1. P. Fishburn, *Utility theory for decision making*, 1970
  2. D. Kreps, *Notes on the theory of choice*, 1988
- Classical presentations of the "neo-classical" theory:
  1. I. Gilboa, *Theory of decision under uncertainty*, 2009
  2. P. Wakker, *Prospect theory*, 2010

Some readings

- Recent surveys and overviews which the tutorial is based upon:

1. I. Gilboa and M. Marinacci, Ambiguity and the Bayesian paradigm, 2013 (in a *Cambridge U. Press* book)
2. L. P. Hansen, Nobel lecture: Uncertainty outside and inside economic models, *J. Political Economy*, 2014
3. L. P. Hansen and M. Marinacci, Ambiguity aversion and model misspecification: An economic perspective, *Stat. Science*, 2016
4. M. Marinacci, Model uncertainty, *J. Europ. Econ. Ass.*, 2015
5. L. Berger et al., Rational policymaking during a pandemic, *PNAS*, 2021