Making Decisions under Model Misspecification

Simone Cerreia-Vioglio, Lars Peter Hansen, Fabio Maccheroni, Massimo Marinacci

TD III

1 luglio 2021, Università di Parma

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
- Environments with uncertainty through the guise of models (e.g., policy making)
- Decision making under model uncertainty

Savage Decision problems

A 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, this quartet takes the *Anscombe-Aumann* form
- We abstract from state misspecification issues (e.g., unforeseen contingencies)

Probability models

- Δ is the set of probability measures on S
- DMs posit a set Q of *structured* (probability) *models* $q \in \Delta$ on states, with a substantive motivation or scientific underpinnings
- Each q describes a possible *DGP*, so it represents *physical uncertainty* (risk)

Probability models

- DMs thus posit a model space Q in addition to the state space S
- When the model space is based on experts' advice, its nonsingleton nature may reflect different advice
- Denti and Pomatto (2019) came up with a behavioral counterpart of Q
- To ease matters, Q is a convex and compact subset of Δ^{σ}

Uncertainty: a taxonomy

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

Uncertainty: a taxonomy

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

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

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

- There is no decision criterion that accounts for model misspecification concerns
- 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

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

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

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

$$V(f) = \min_{q \in Q} \int u(f)\, dq$$

  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 Q is a singleton $\{q\}$, so no model ambiguity, we have the multiplier criterion

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

- 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

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

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

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

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

Box and all that

- Structured models may be incorrect, yet useful as Box (1979) famously remarked
- Formally, betting behavior must be *consistent* with datum Q, i.e.,

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

- 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

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

is bet-consistent

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

Misspecification neutrality

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

$$\int u(f)\,dq \geq \int u(g)\,dq \quad \forall q \in Q \Longrightarrow f \succsim g$$

  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

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

  is misspecification neutral iff it is represented by the maxmin criterion

$$V(f) = \min_{q \in Q} \int u(f)\, dq$$

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

A tale of two preferences

- We axiomatize this criterion 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)

Rational preference

A.1 A preference $\succsim$ is (*subjectively*) *rational* if it is:

(a) complete

(b) risk independent: if $x, y, z \in X$ and $\alpha \in (0,1)$, then $x \sim y$ implies $\alpha x + (1-\alpha) z \sim \alpha y + (1-\alpha) z$

- Risk independence implies that $\succsim$ on X (e.g., on lotteries) is represented by an affine (e.g., expected) utility function $u : X \to \mathbb{R}$

Dominance relation

A.2 A preference $\succsim^*$ on $\mathcal{F}$ is a *dominance relation* (or is *objectively rational*) if it is:

(a) c-complete: if $x, y \in X$, then $x \succsim^* y$ or $y \succsim^* x$

(b) completess: when Q is a singleton, $f \succsim^* g$ or $g \succsim^* f$ for all $f, g \in \mathcal{F}$

(c) weak c-independent: if $f, g \in \mathcal{F}$, $x, y \in X$, and $\alpha \in (0, 1)$,

$$\alpha f + (1-\alpha)x \succsim^* \alpha g + (1-\alpha)x \Rightarrow \alpha f + (1-\alpha)y \succsim^* \alpha g + (1-\alpha)y$$

(d) convex: if $f, g, h \in \mathcal{F}$ and $\alpha \in (0, 1)$,

$$f \succsim^* h \text{ and } g \succsim^* h \Rightarrow \alpha f + (1-\alpha)\, g \succsim^* h$$

Dominance relation

- Model ambiguity (i.e., a nonsingleton Q) underlies the incompleteness of $\succsim^*$
- Because of model misspecification, $\succsim^*$ satisfies only a weak form of independence (even when Q is a singleton)

Consistency

A.3 Preferences $\succsim^*$ and $\succsim$ are *consistent* if

$$f \succsim^* g \Longrightarrow f \succsim g$$

for all $f, g \in \mathcal{F}$

Caution

- **A.4** Preferences $\succsim^*$ and $\succsim$ satisfy *caution* if

$$f \not\succsim^* x \Longrightarrow x \succsim f$$

for all $x \in X$ and all $f \in \mathcal{F}$

- DM opts, by default, for a sure alternative x over an uncertain one f, unless the dominance relation says otherwise

Q-Coherent (objective)

- We write $f \overset{Q}{=} g$ when $q\left(f = g\right) = 1$ for all $q \in Q$
- **A.5** A dominance relation $\succsim^*$ on $\mathcal{F}$ is (*objectively*) *Q-coherent* if
$$f \overset{Q}{=} g \implies f \sim^* g$$

Q-Coherent (subjective)

- Given a model $p \in \Delta$, define a consequence $x_f^p \in X$ for each act f via the equality

$$u(x_f^p) = \int u\left(f\right) dp$$

- We can interpret x_f^p as the "certainty" equivalent of act f if p were the correct model

Q-Coherent (subjective)

- **A.6** A rational preference $\succsim$ is (*subjectively*) *Q-coherent* if, for all $f \in \mathcal{F}$ and all $x \in X$, we have

$$x \succ^* x_f^p \Longrightarrow x \succ f$$

  if and only if $p \in Q$

- A probability p is a structural model when DMs take it seriously, i.e., they never choose acts that would be strictly dominated if p were correct

Divergences

Given a (non-empty) subset Q of Δ, a function $c : \Delta \times Q \longrightarrow [0, \infty]$ is a *divergence* (*for the set* Q) if

(i) the sections $c_q : \Delta \longrightarrow [0, \infty]$ are grounded, lsc and convex for each $q \in Q$

(ii) the function $c_Q : \Delta \longrightarrow [0, \infty]$ defined by

$$c_Q(p) = \min_{q \in Q} c(\cdot, q)$$

is well defined, grounded, lsc and convex

(iii) $c_Q^{-1}(0) = Q$, that is, $c_Q(p) = 0$ if and only if $p \in Q$

Statistical distances

- A divergence c that satisfies the distance property

$$c\left(p,q\right)=0\iff p=q$$

  is called *statistical distance*

- $c_Q\left(p\right)$ is now an Hausdoff-type statistical distance between p and Q

Representation

THM Let $(S, \Sigma, X, Q, \succsim^*, \succsim)$ be a two-preference classical decision environment. The following statements are equivalent:

(i) $\succsim^*$ is a unbdd dominance relation and $\succsim$ is a rational preference that are both Q-coherent and jointly satisfy consistency and caution;

Representation

(ii) there exist an onto affine function $u : X \to \mathbb{R}$ and a divergence $c : \Delta \times Q \to [0, \infty]$, with $\mathrm{dom}\, c_Q \subseteq \Delta(Q)$, such that, for all acts $f, g \in \mathcal{F}$,

$$f \succsim^{*} g \Leftrightarrow \min_{p \in \Delta} \left[\int u(f)\, dp + c(p, q) \right] \geq \min_{p \in \Delta} \left[\int u(g)\, dp + c(p, q) \right]$$

for all $q \in Q$, and

$$f \succsim g \Leftrightarrow \min_{p \in \Delta} \left[\int u(f)\, dp + \min_{q \in Q} c(p, q) \right] \geq \min_{p \in \Delta} \left[\int u(g)\, dp + \min_{q \in Q} c(p, q) \right]$$

If, in addition, c is uniquely null, then $c : \Delta \times Q \to [0, \infty]$ can be chosen to be a statistical distance.

Representation

- This result identifies the main preferential assumptions underlying a representation

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

where $c : \Delta \times Q \to [0, \infty]$ is a statistical distance

- Let $\succsim_1$ be $\succsim_2$ represented by this criterion with same u

- $\succsim_1$ is more uncertainty averse than $\succsim_2$ iff

$$\min_{q \in Q} c_1(\cdot, q) \leq \min_{q \in Q} c_2(\cdot, q)$$

Representation

- The function

$$p \mapsto \min_{q \in Q} c\left(p, q\right)$$

as an index of aversion to model misspecification, for short, a *misspecification index*

- The lower is this index, the higher is the fear of misspecification

Varying structured information

- So far, a given set Q of structured models
- We should write
$$\left(S, \Sigma, X, \succsim_Q^*, \succsim_Q\right)$$
with the dependence of preferences on Q highlighted
- Yet, decision environments may share common state and consequence spaces, but differ on the posited sets of structured models because of different information that DMs may have
- Need of decision criteria that, across such environments, are consistent

Varying structured information

- Let $\mathcal{Q}$ be the family of convex and compact subsets of Δ^σ
- Consider a family

$$\left\{\left(S, \Sigma, X, \succsim^*_Q, \succsim_Q\right)\right\}_{Q\in\mathcal{Q}}$$

  of decision environments that differ in the set Q of posited models
- Introduce axioms that connect these environments

Varying structured information

A.7 The family $\{\succsim^*_Q\}_{Q\in\mathcal{Q}}$ is *monotone* (*in model ambiguity*) if $Q' \subseteq Q$ implies, for all $f, g \in \mathcal{F}$,

$$f \succsim^*_Q g \Longrightarrow f \succsim^*_{Q'} g$$

- If the "structured" information underlying a set Q is good enough for the DM to establish that an act dominates another one, a better information which decreases model ambiguity can only confirm such judgement

- Its reversal would be at variance with the objective rationality spirit of the dominance relation

Varying structured information

A.8 The family $\{\succsim^*_Q\}_{Q \in \mathcal{Q}}$ is *Q-separable* if, for each $f, g \in \mathcal{F}$ and $x \in X$,

$$\forall q \in Q, f \succsim^*_q g \implies f \succsim^*_Q g$$

- If, without model ambiguity, an act unanimously dominates another one according to each model q in Q, so does under model ambiguity when Q is the set of structured models

- No model in Q supports a reversal of this dominance judgement

Varying structured information

- Previous axioms imply that $\succsim_Q^*$ agree on X
- We thus just write $\succsim^*$

Varying structured information

- Denote by $x_{f,q}$ the certainty equivalent of act f for preference $\succsim_q^*$

A.7 *Model hybridization aversion*. Given any $q, q' \in \Delta^\sigma$,

$$\lambda x_{f,q} + (1-\lambda) x_{f,q'} \succsim^* x_{f,\lambda q + (1-\lambda) q'}$$

for all $\lambda \in (0,1)$ and all $f \in \mathcal{F}$

- DM dislikes, *ceteris paribus*, facing a hybrid structured model $\lambda q + (1-\lambda) q'$ that, by mixing two structured models q and q', could only have a less substantive motivation

Varying structured information

THM Let

$$\left\{\left(S,\Sigma,X,\succsim_Q^*,\succsim_Q\right)\right\}_{Q\in\mathcal{Q}}$$

be a family of two-preference classical decision environments. The following statements are equivalent:

(i) $\left\{\succsim_Q^*\right\}_{Q\in\mathcal{Q}}$ is monotone, Q-separable and, for each $Q\in\mathcal{Q}$, the preferences $\succsim_Q^*$ and $\succsim_Q$ satisfy the hypotheses of Theorem 1;

Varying structured information

(ii) there exist an onto affine function $u : X \longrightarrow \mathbb{R}$ and a lsc and convex statistical distance $c : \Delta \times \Delta^{\sigma} \longrightarrow [0, \infty]$, with $\mathrm{dom}\, c_Q \subseteq \Delta(Q)$ for each $Q \in \mathcal{Q}$, such that, for all acts $f, g \in \mathcal{F}$,

$$f \succsim_Q^* g \Leftrightarrow \min_{p \in \Delta} \left[\int u(f)\, dp + c(p, q) \right] \geq \min_{p \in \Delta} \left[\int u(g)\, dp + c(p, q) \right]$$

for all $q \in Q$, and

$$f \succsim_Q g \Leftrightarrow \min_{p \in \Delta} \left[\int u(f)\, dp + \min_{q \in Q} c(p, q) \right] \geq \min_{p \in \Delta} \left[\int u(g)\, dp + \min_{q \in Q} c(p, q) \right]$$

Moreover, u is unique up to a positive affine transformation and, given u, c is unique

To be continued

- Bayesian analysis
- Dynamic analysis
- Applications