

Calibration and Subjective Confidence Intervals

Answers to the questions at the end of the last class:

Stylized fact: most people's confidence intervals are "too small."

- How is this related to subjective expected utility and its axioms?
- Note there is a really cheap way of making sure that 90% of your 90% confidence intervals contain the truth.
- Alpert and Raiffa [1982] (a chapter of *Judgement Under Uncertainty*) :46% of quantities outside of 98% confidence intervals, 40% outside of 99.8% confidence intervals.
- Biais, Hilton, Mazurier, and Pouget, *REStud* [2005]: less over-confident traders do better in an experimental asset market.

- Many physicians are overconfident, especially residents (see e.g. survey by Berner and Graber *Am. J. Medicine* [2008]: *“Residents... were more confident about the correctness of their diagnoses, but they were less accurate than the attending physicians. [In] a cohort of 95 board-certified radiologists... the confidence level of the worst performers was actually higher than that of the top performers.”*)
- Feedback on accuracy of confidence intervals leads to better calibration.
- As does giving advice: “spread your extreme quantiles, you probably shouldn’t be so certain of your beliefs...”
- Calibration of subjective confidence intervals is based on self-knowledge, and can be applied to non-random events.

- Can also look at calibrated forecasts of things that happen repeatedly: A weather forecaster is *calibrated* if on those occasions when the forecaster announced probability p of rain, it actually rained fraction p of the time.
- U.S. weather forecasters are pretty well calibrated. (*though published reports may bias upwards the probability of precipitation*)
- Calibrated forecasts needn't be useful as miscalibrated ones. For example forecaster who correctly predicts the average chance of rain over the year can miss large seasonal effects.
- And with a large enough data set a forecaster with no knowledge at all can ensure that she is calibrated in the long run. (Foster and Vohra, *Biometrika* [1998]).
- But it still may seem like a good thing to try for.

The Ellsberg Paradox

- 90 Balls in an urn composed of red, black and green. 30 of the balls are definitely red, while the other 60 could be black or green.

- $S = \{R, G, B\}$, 4 acts f, g, f', g' that pay \$ x or 0.

	R	G	B
f	x	0	0
• g	0	x	0
f'	x	0	x
g'	0	x	x

- Preferences $f \succ g$ and $f' \prec g'$ not consistent with subjective EU: $f \succ g$ implies $\Pr(R) > \Pr(G)$ while $f' \prec g'$ implies $\Pr(R) < \Pr(G)$.

- *Reason:* if the preferences satisfy Independence, the above choices would require

$$\frac{1}{2}f + \frac{1}{2}g' \succ \frac{1}{2}g + \frac{1}{2}g' \succ \frac{1}{2}g + \frac{1}{2}f' .$$

- But these 2 lotteries are the same, they both give the constant act with .5 chance of x in every state. (*what implicit assumption am I using here?*)
- Related thought experiment of Keynes: suppose urn 1 has 50 green balls and 50 yellow balls. Urn 2 has 100 balls, each either green or yellow, but the proportion is unknown. Keynes [1921] argued that it would be reasonable to be indifferent between betting on a green or yellow draw from urn 1, and prefer both bets to either betting on green or yellow from urn 2.
- But this too contradicts subjective EU.

- Say that someone is “ambiguity averse” if they prefer the objective lottery to the subjective ones.
- Similar violations of independence have been observed in many experiments, see Camerer and Weber *J Risk Uncertainty* [1992], Trautmann and Kuilen, Ch. 3 of *Handbook of Judgment and Decision Making* [2015].
- In response, various models of *ambiguity averse* preferences relax the independence axiom.

- *Maxmin preferences* (Gilboa-Schmeidler *J Math Econ* [1989]): start from the Anscombe-Aumann setup. As before, take S finite and let $\mathcal{F} = (\Delta(Z))^S$ be the space of objective lotteries over prizes, one lottery per state.
- Drop the independence axiom, replace it with the next two conditions.

Certainty Independence:

For all constant acts $\vec{\ell}$ and all $f, g \in \mathcal{F}$,

$f \succsim g$ iff $\alpha f + (1-\alpha)\vec{\ell} \succsim \alpha g + (1-\alpha)\vec{\ell}$ for all $\alpha \in (0,1)$.

- Weaker than independence, which says $f \succsim g$ iff $\alpha f + (1-\alpha)h \succsim \alpha g + (1-\alpha)h$ for any act h .
- *Reasons this might fail...?*

Uncertainty Aversion: Whenever $f \sim g$, $\alpha f + (1-\alpha)g \succsim f$ $\alpha \in (0,1)$.
(could also call this a preference for hedging).

- To understand this name suppose there are two states R and B .
- act f pays 1 in R and 0 in B ; act g pays 0 in R and 1 in B .
- Then $\alpha f + (1-\alpha)g$ pays $(\alpha, (1-\alpha))$ so “less uncertain,” $\alpha = 1/2$ pays $EU=1/2$ regardless of which state occurs.

Theorem (Gilboa and Schmeidler) Complete transitive preferences on \mathcal{F} satisfy mixture continuity, monotonicity, non-triviality, certainty independence, and uncertainty aversion iff there is a linear, non-constant $u : \Delta(Z) \rightarrow \mathbb{R}$ and a non-empty, closed, convex set of distributions $C \in \Delta(S)$ s.t.

$$V(f) := \min_{p \in C} \sum_s p(s)u(f(s)) .$$

Moreover, C is identified uniquely, and u is unique up to affine transformations.

In this representation, C can be viewed as the priors the agent thinks are possible, and she tries to maximize her worst possible payoff: she acts as if Nature were out to get her.

Maxmin preference: maximize $\min_{p \in C} \sum_s p(s)u(f(s))$.

Examples:

- $C = \{p\}$ for some single p is subjective expected utility.
- $C = \Delta(S)$ is very pessimistic: evaluate every act by its worst possible result.

- In the 3-color Ellsberg urn described earlier, one candidate for C is $C = \{p \in \Delta(S) : p(R) = 1/3\}$.
- This allows the Ellsberg choices: Set $u(x) = 1, u(0) = 0$.

	R	G	B
f	1	0	0
g	0	1	0
f'	1	0	1
g'	0	1	1

Then

$$\min_{p \in C} \sum_s p(s)u(f(s)) = 1/3 > 0 = \min_{p \in C} \sum_s p(s)u(g(s))$$

$$\min_{p \in C} \sum_s p(s)u(f'(s)) = 1/3 < 2/3 = \min_{p \in C} \sum_s p(s)u(g'(s)).$$

Idea of proof of the Gilboa-Schmeidler representation theorem.

Start out following our proof of the Anscombe-Aumann theorem:

1. Show the induced preference on $\Delta(Z)$ is represented by a non-trivial utility function u that is linear in the objective probabilities. Normalize the range of u to be $U = [-1,1]$.
2. Define *utility acts* $\tau : S \rightarrow U$ and define \succsim^* on U^S by $\tau \succsim^* \tau'$ iff $f \succsim g$ for some $f, g \in \mathcal{F}$ s.t. $\tau = u \circ f$ and $\tau' = u \circ g$.
3. For each utility act τ , let $I(\tau)$ be the number s.t. $\overline{I(\tau)} \sim \tau$.
(as before need to show this is well defined.)

In A-A the indifference function I was additive, now it need not be, as the a agent may strictly prefer a mixture of two acts (to “hedge.”)

Instead, we can use the assumptions on preferences to show I is:

-monotone (*same definition as in A-A: If $\overrightarrow{f(s)} \succsim \overrightarrow{g(s)} \forall s$ then $f \succsim g$*)

-normalized (for any u , $I(\vec{u}) = u$),

-positively homogeneous (*same proofs for these as in AA, but we use certainty independence in place of independence*)

-“translation invariant” (for any $\tau \in U^S, k \in \mathbb{R}, I(\tau + \vec{k}) = I(\tau) + k$)

(*from certainty-independence and homogeneity*).

-concave (*from uncertainty aversion*)

Use translation invariance and monotonicity to show we can extend I to all of R^S (*Gilboa-Schmeidler let S be an arbitrary set, and work on the space of all bounded measurable functions*)

Proof of concavity: from homogeneity and extension of I , it is enough to show

$$I\left(\frac{1}{2}\tau + \frac{1}{2}\tau'\right) \geq \frac{1}{2}I(\tau) + \frac{1}{2}I(\tau') \text{ for any } \tau, \tau' \in U^S .$$

(for other combinations, rewrite $\alpha\tau + (1-\alpha)\tau' = \frac{1}{2}v + \frac{1}{2}v'$ for $v = 2\alpha\tau, v' = 2(1-\alpha)\tau$.)

- Fix f, g , s.t. $u \circ f = \tau, u \circ g = \tau'$.

- If $I(\tau) = I(\tau')$, uncertainty aversion implies $\frac{1}{2}f + \frac{1}{2}g \succsim f$,

$$\text{so } I\left(\frac{1}{2}\tau + \frac{1}{2}\tau'\right) \geq I(\tau) = \frac{1}{2}I(\tau) + \frac{1}{2}I(\tau').$$

- If $I(\tau) > I(\tau')$ define $\delta = I(\tau) - I(\tau')$ and $\tau'' = \tau' + \bar{\delta}$.

Then $I(\tau'') = I(\tau') + \delta$ (from translation invariance) so

$$\begin{aligned} I\left(\frac{1}{2}\tau + \frac{1}{2}\tau'\right) + \frac{1}{2}\delta &= I\left(\frac{1}{2}\tau + \frac{1}{2}\tau''\right) \\ &\geq \frac{1}{2}I(\tau) + \frac{1}{2}I(\tau'') = \frac{1}{2}I(\tau) + \frac{1}{2}I(\tau') + \frac{1}{2}\delta \end{aligned}$$

So I is concave.

The representation theorem now follows from a result in convex analysis: functions I with the properties we established above can be identified with a unique convex set $C \in \Delta(S)$ s.t. $I(\tau) := \min_{p \in C} \sum_s \tau(s) p(s)$ - this set C is the set of priors in the minmax representation.

This uses the idea of *support functions*:

Definition: For closed convex and non-empty $C \subseteq \Delta(S)$, the *support function* $R_C : \mathbb{R}^S \rightarrow \mathbb{R}$ is

$$R_C(\tau) := \inf_{p \in C} \sum_s p(s) \tau(s).$$

If p is a vector of inputs, C is the set of all inputs that produce a given output, and τ is a price vector, this is the problem of finding the least-cost way of producing a given output. (*if we let the firm both buy input and sell output, and consider a supremum instead of an infimum, we compute the profit of a given production plan.*)

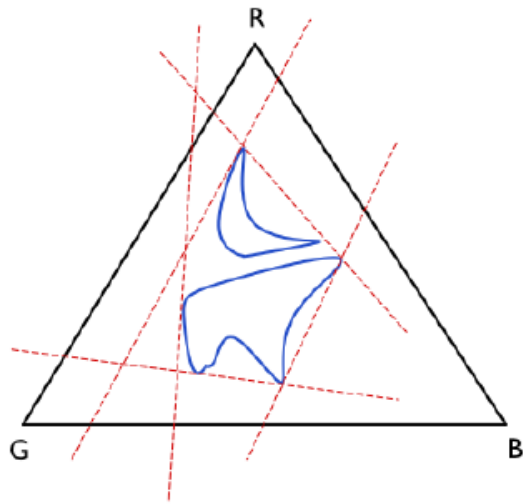


Figure 10.5. Finding points that are minimizers for some acts.
 C is indistinguishable from its convex hull.

The support function of any convex set is positive homogeneous and translation invariant.

Because I is concave, monotone, and normalized, we can replace the inf with a min, and uniquely identify C . (see e.g. Rockafellar [1997] Cor. 13.2.2)

- Characterizing I as the support function of set is an example of *conjugate duality*.
- *Support functions also used in mechanism design, see Goeree and Kushnir GEB [2016] and their working paper “Geometric Approach to Mechanism Design.”*
- With minmax preferences, a probability distribution p is either in or out of the support.
- More general representations allow for different non-zero weights on different beliefs.

- *Variational preferences* (Maccheroni, Marinacci, and Rustichini *Ema* [2006]):

$$V(f) = \min_{p \in \Delta(S)} \sum_s p(s)u(f(s)) + c(p) ,$$

where $c : \Delta(S) \rightarrow [0, \infty]$, $c(p) = 0$ for at least one p , and is convex and lower semi-continuous.

- Note that this nests minmax EU: set $c = 0$ or ∞ for p in or out of C , but more general, so corresponds to weaker conditions on the function l .
- Variational preferences also nests the “multiplier preferences” of Hansen and Sargent *AER* [2001], where $c = \theta R(p \parallel q)$ is a constant times the relative entropy between p and a fixed function q :

$$R(p \parallel q) = \sum_s p(s) \ln \left(\frac{p(s)}{q(s)} \right), p \ll q; \text{ else } R = \infty.$$

- Here q represents the agent’s “best guess” of the distribution. ($p \ll q$ means p is absolutely continuous w.r.t q so the ratio is well defined.)

- HS's motivation comes from robust control theory, where q is a best estimate given a model the agent thinks may be mis-specified; Bayesian posterior converges to a point mass on the distribution that minimizes the divergence.
- As θ grows, the agent is more focused on p 's that are "very near" q .

Variational preferences are characterized by almost the same axioms as for maxmin, with “weak certainty independence” (*defined at the end of these notes*) replacing independence.

The proof then uses a more general result about convex duality:

$V(f) = \min_{p \in \Delta(S)} \sum_s p(s)u(f(s)) + c(p)$ is the *Legendre conjugate* of the function c .

MMR show that the indifference function V satisfies the conditions for being the Legendre conjugate of a convex function c and that this function is uniquely determined.

Ghili and Klibanoff [2017] show that all variational preferences (including maxmin) that fit the Ellsberg paradox violate the following “*monotonicity in mixtures*” axiom:

For all acts f, f', g s.t. $\overrightarrow{f'(s)} \succsim \overrightarrow{f(s)} \quad \forall s$, and all $1 \geq \alpha' \geq \alpha \geq 0$,

$$\alpha' f + (1 - \alpha')g \succsim \alpha f + (1 - \alpha)g \rightarrow \alpha' f' + (1 - \alpha')g \succsim \alpha f' + (1 - \alpha)g,$$

and the strict-preference version,

$$\alpha' f + (1 - \alpha')g \succ \alpha f + (1 - \alpha)g \rightarrow \alpha' f' + (1 - \alpha')g \succ \alpha f' + (1 - \alpha)g.$$

This says that if the decision maker likes moving probability weight from g to f , she also likes moving it from g to the dominating act f' . *Comments?*

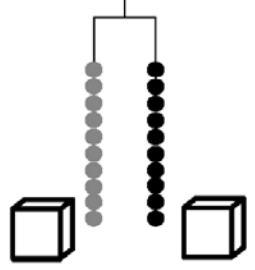
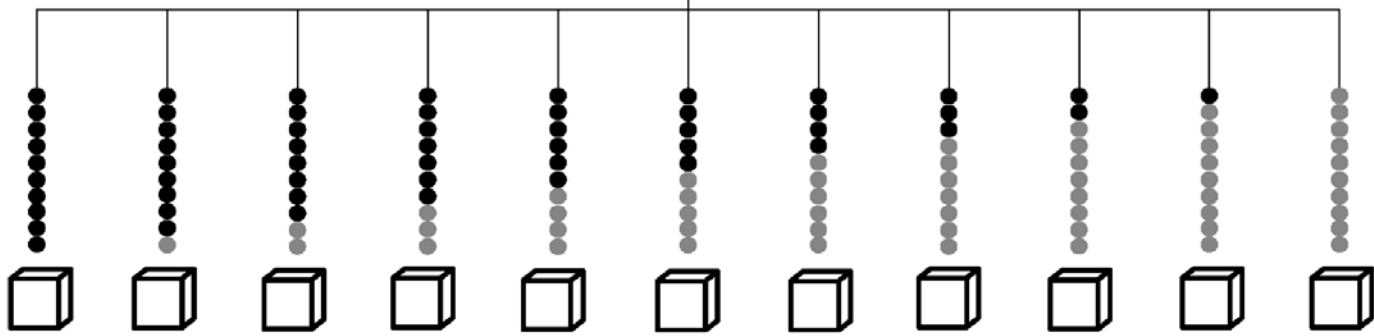
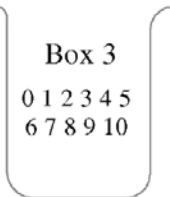
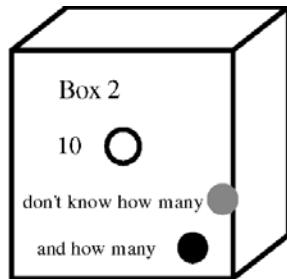
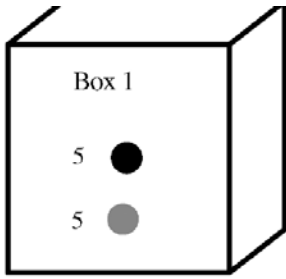
Note that monotonicity in mixtures limits the agent’s preference for hedging...

Some forms of “second order variational preferences” (Klibanoff, Marinaci, and Mukerji *Ema* [2005], Stzralecki section 10.9) can accommodate both Ellsberg and this monotonicity condition.

Halevy *Ema* [2007]: ambiguity-aversion is highly correlated with not reducing *objective* compound lotteries

4 “urns”: (*the experiment called these “boxes,” implemented via computer*)

- *urn 1*: 5 red balls and 5 black balls.
 - *urn 2*: 10 balls, each either red or black, composition unspecified.
 - *urn 3*: Number of red chosen uniformly from [0,10].
 - *urn 4*: (1/2,1/2) randomization between all red and all black.
- If subjects reduce compound lotteries, urns 1, 3, and 4 are equivalent; urn 2 is “ambiguous.”
 - Subjects bet on red or black from each urn.
 - Then given chance to sell their bets using the Becker DeGroot Marshak (BDM) mechanism.



- The argument that “bid your value” is optimal in BDM relies on reduction of compound lotteries and the independence axiom, so may not work well for ambiguity-averse subjects.
- But anyone who reduces compound lotteries should have the same certain equivalent for bets on 1, 3, and 4.
- Two rounds, second “robustness” round had higher stakes, slightly different instructions.

“Subjects who reduced compound lotteries were almost always ambiguity neutral, and most subjects who were ambiguity neutral reduced compound lotteries appropriately (15–20% of the subjects). The remainder of the subjects exhibited violations of ROCL (reduction of compound lotteries) and ambiguity aversion, but there is no unique theory that can accommodate the different choice patterns in the population. The population is heterogeneous and two choice patterns, which account for approximately 70% of all subjects, emerge... that currently there is no unique theoretical model that universally captures ambiguity preferences.”

- The correlation of ambiguity aversion and failure to reduce compound lotteries was replicated by Chew, Miao, and Zhong *Ema* [2017] and Baillon and Bleichcrodt *AEJ:Micro* [2017].
- In some situations, people seem to be ambiguity-seeking instead of ambiguity-averse.
- For example urn 1 has 10 colors, 10% each; urn 2 has same 10 colors, unknown proportions. Would you rather “bet on red” in urn 1 or urn 2? Many subjects prefer urn 2.
- For moderate and high probability events in the “gains domain,” many people prefer objective lotteries and so are “ambiguity averse.”
- However, the prevalence and magnitude of ambiguity aversion depends on the elicitation method: people are less ambiguity averse when asked to choose an objective “probability equivalent” than when asked for certain equivalents.

- And there is some evidence that people have different ambiguity attitudes in the loss domain.
- Baillon and Bleichcrodt *AEJ:Micro* [2017] observed a “*fourfold pattern of ambiguity attitudes: ambiguity aversion for likely gains and unlikely losses and ambiguity seeking for unlikely gains and likely losses.*”
- For more on this see the “Ambiguity Attitudes” survey by Trautmann and van de Kuilen in *Handbook of Judgment and Decision Making* [2015].
- So the question of how to model ambiguity aversion is far from settled...

Reading for next class: Strzalecki ch. 2, 3.1-3.3

Appendix: *Weak Certainty Independence*

- Variational preferences imply the following relaxation of certain-independence:

For all $f, g \in \mathcal{F}$, $\alpha \in (0,1)$, and constant acts $\vec{\ell}, \vec{\ell}'$,

$$\alpha f + (1-\alpha)\vec{\ell} \succsim \alpha g + (1-\alpha)\vec{\ell} \Leftrightarrow \alpha f + (1-\alpha)\vec{\ell}' \succsim \alpha g + (1-\alpha)\vec{\ell}'$$

- Weaker than certainty independence ($f \succsim g$ iff $\alpha f + (1-\alpha)\vec{\ell} \succsim \alpha g + (1-\alpha)\vec{\ell}$)

which is equivalent to for all $f, g \in \mathcal{F}$, $\alpha, \beta \in (0,1]$, and constant acts $\vec{\ell}, \vec{\ell}'$,

$$\alpha f + (1-\alpha)\vec{\ell} \succsim \alpha g + (1-\alpha)\vec{\ell} \text{ iff } \beta f + (1-\beta)\vec{\ell}' \succsim \beta g + (1-\beta)\vec{\ell}'.$$

- Setting $\alpha = \beta$ here gives weak certainty dependence and translation invariance.