What are we weighting for?

A mechanistic model for probability weighting

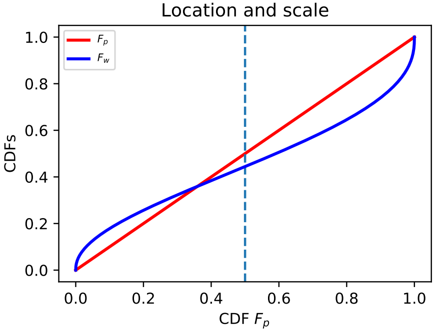
Ole Peters Alexander Adamou Yonatan Berman Mark Kirstein

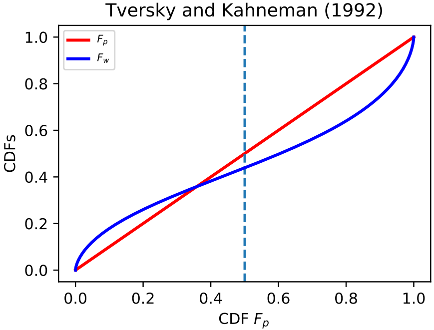
✉ m.kirstein@lml.org.uk
@nonergodicMark

D-TEA 2020, 16 June 2020

Max Planck Institute for Mathematics in the Sciences
UNIVERSITÄT LEIPZIG Mathematisches Institut
LONDON MATHEMATICAL LABORATORY

Main results

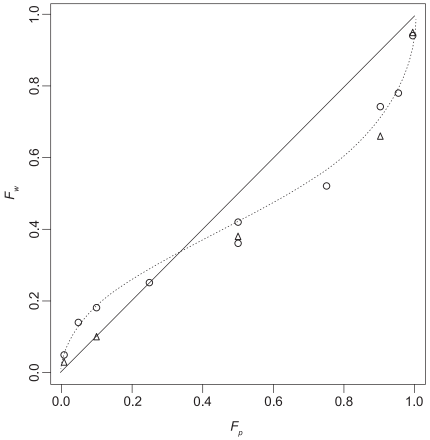




1. generic inverse-S shape can be explained by difference in uncertainty
2. process of estimation of this uncertainty generates inverse-S shape

PW K&T 1979

Defining Probability Weighting (PW)



(Tversky and Kahneman 1992, p. 310, Fig. 1, relabelled axes)

- overestimation of rare events $\rightarrow$ underestimation of common events
- stable empirical pattern: inverse-S shape

Received wisdom:

- PW = maladaptive irrational cognitive bias

In search of a mechanism

$\hookrightarrow$ How does this pattern emerge?

$\hookrightarrow$ Can we derive the functional form (rather than merely fitting some function)?

Set up : A Thought Experiment

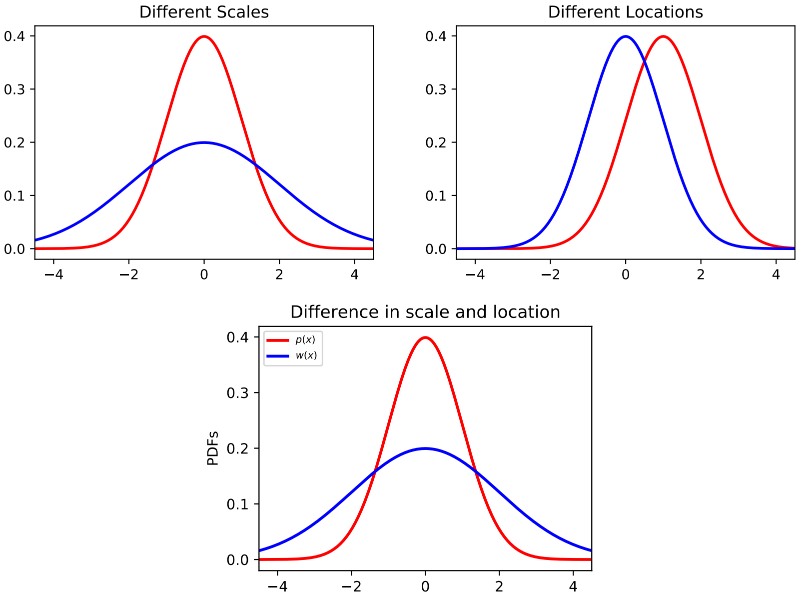
Disinterested Observer (DO)

DO has a model of the random variable X, *e.g.* payout of a gamble
probabilities $p(x)$
CDF $F_p(x)$

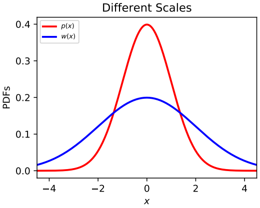
Decision Maker (DM)

DM has a different model of the same random variable X with greater uncertainty
decision weights $w(x)$
CDF $F_w(x)$

Types of Different Uncertainties



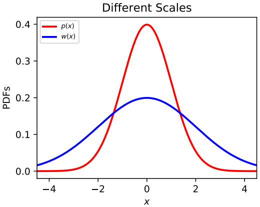
The Simplest Case : Different Scales



Numerical procedure applies to arbitrary distributions:

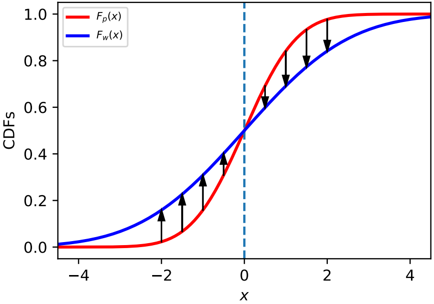
1. construct a list of values for the CDF assumed by the DO, $F_p(x)$
2. construct a list of values for the CDF assumed by the DM, $F_w(x)$
3. plot $F_w(x)$ *vs.* $F_p(x)$

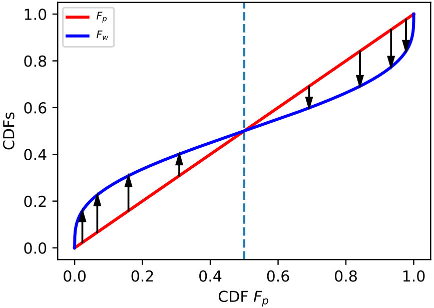
The Simplest Case : Different Scales



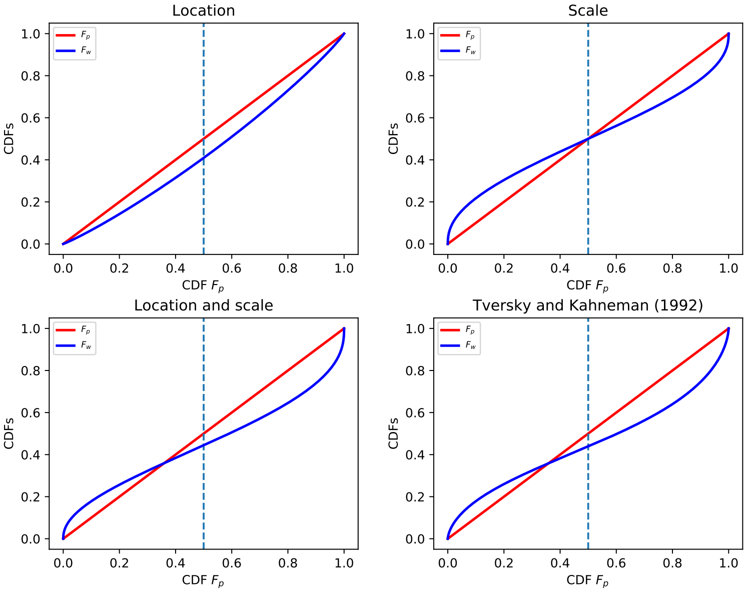
Numerical procedure applies to arbitrary distributions:

1. construct a list of values for the CDF assumed by the DO, $F_p(x)$
2. construct a list of values for the CDF assumed by the DM, $F_w(x)$
3. plot $F_w(x)$ *vs.* $F_p(x)$





Applying the Procedure to the Uncertainty Types

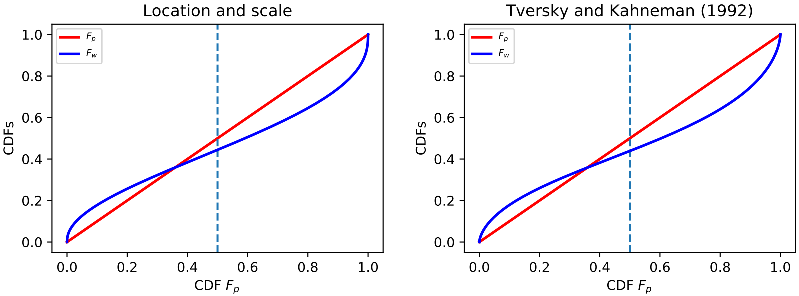


Interim conclusion

- greater scale reproduces inverse-S shape
- differences in location and scale reproduce asymmetric inverse-S shape
- inverse-S shape arises for all unimodal distributions
- Probability Weighting is the effect of a difference in uncertainty

Job done. Thank you for your attention ;)

Functional Forms



Asking the Ergodicity Question

DO's concern

What happens on average to the ensemble of subjects?

$\neq$

DM's concern

What happens to me on average over finite time?

Extra Uncertainty is Part of DM's Inference Problem I

DM's adaptive rationality: err on the side of caution:

- DM has no control over the experiment,
- DM's incomplete comprehension of the experiment/decision problem,
- DM needs to trust the DO
- uncertain outcome is consequential only to the DM,
- ...

Extra Uncertainty is Part of DM's Inference Problem II

- "probability" is polysemous (Gigerenzer 1991, 2018; Hertwig and Gigerenzer 1999)
- probabilities are not observable, but
- DM observes counts of (rare) events along his life trajectory through time

$\hookrightarrow$ **DM's inference problem:** estimate probability $p(x)$ from counts

Nature of Inference for Rare Events

Rare Event

- $p(x) = 0.0001$
- 10000 observations
- $\sim$ 99.5% of such time series will contain 0 or 1 events
- Naïve estimation: $\hat{p}(x) = 0$ or $\hat{p}(x) = 0.0001$

$\hookrightarrow$ either impossible or ten times (over)estimation

Common Event

- $p(x) = 0.1$
- 10000 observations
- $\sim$ 99.5% of time series would contain between 50 and 150 events,

$\hookrightarrow$ much smaller relative error in $\hat{p}(x)$

$\hookrightarrow$ the smaller $p(x)$ the smaller the count of it in a finite time series
$\hookrightarrow$ the bigger the relative estimation error

Relative Estimation Error is Larger for Rarer Events

Asymptotic probability	Most likely count	Standard error in count	Standard error in probability	Relative error in probability
0.1	1000	32	0.003	3%
0.01	100	10	0.001	10%
0.001	10	3	0.0003	30%
0.0001	1	1	0.0001	100%

Table: $T = 10\,000$, assuming Poisson statistics, relative estimation errors $\sim {}^{1}/\sqrt{\text{count}}$

Estimation of the decision weights

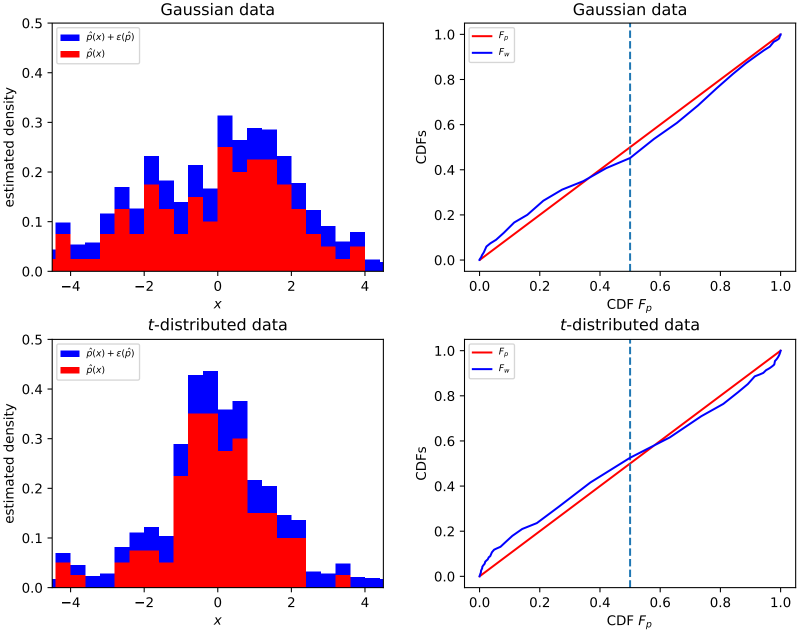
Using the count $n(x)$ to form the best estimate and add to it the uncertainty about best estimate

$$w(x) \approx \frac{n(x)}{T\delta x} \pm \frac{\sqrt{n(x)}}{T\delta x} \tag{1}$$
$$w(x) \approx \hat{p}(x) \pm \varepsilon\left[\hat{p}(x)\right] \tag{2}$$

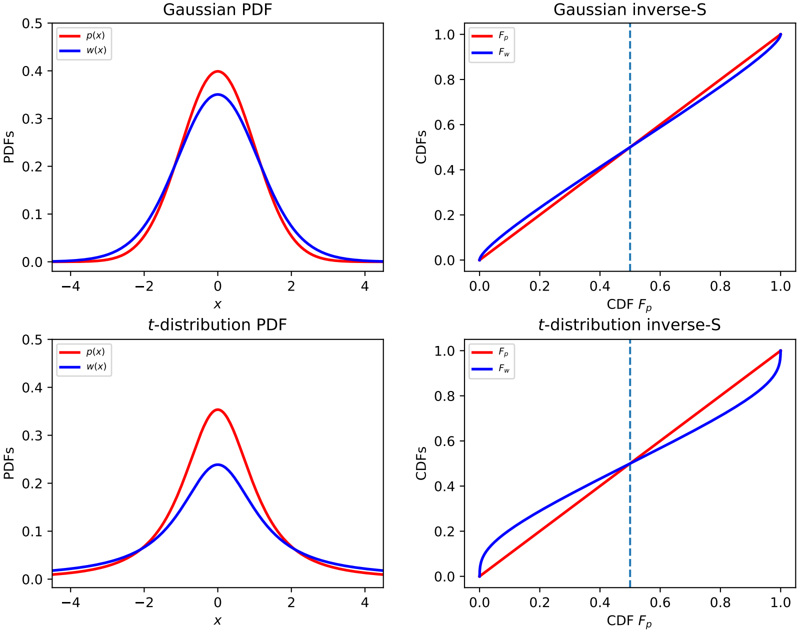
with the standard error expressed in terms of the estimate itself

$$\varepsilon\left[\hat{p}(x)\right] \equiv \frac{\sqrt{n(x)}}{T\delta x} = \sqrt{\frac{\hat{p}(x)}{T\delta x}} \tag{3}$$
$$\lim_{T\to\infty} w(x) \to p(x) \tag{4}$$

Mark Kirstein



Mark Kirstein



Conclusion

Ergodicity Economics explains probability weighting

- inverse-S shape as a neutral indicator of a difference in opinion
- reported observations are consistent with DMs extra uncertainty
- relative uncertainty arises out of the situation of the DM over time
- reproduce the right type of uncertainty, *i.e.* relative errors are larger for rare events

$\hookrightarrow$ Probability weighting is rational cautious behaviour under uncertainty over time

- See full paper at `bit.ly/lml-pw-r1`
- links to play with the code are inside

Thank you for your attention!

I'm looking forward to the discussion
Comments & questions are very welcome, here or to

✉ m.kirstein@lml.org.uk
@nonergodicMark

Submit an open peer review to this paper on
bit.ly/lml-pw-r1

Mark Kirstein

BACK UP

Probability Weighting as an Estimation Issue

"It is important to distinguish overweighting, which refers to a property of decision weights, from the overestimation that is commonly found in the assessment of the probability of rare events. [...] In many real-life situations, overestimation and overweighting may both operate to increase the impact of rare events." (Kahneman and Tversky 1979, p. 281)

$\hookrightarrow$ distinguish between

- uncertainty estimation and
- "weighting"

we analyse the former and find very good agreement with the empirical inverse-S pattern

$\hookrightarrow$ How big is the residual "probability weighting" after accounting for uncertainty estimation?

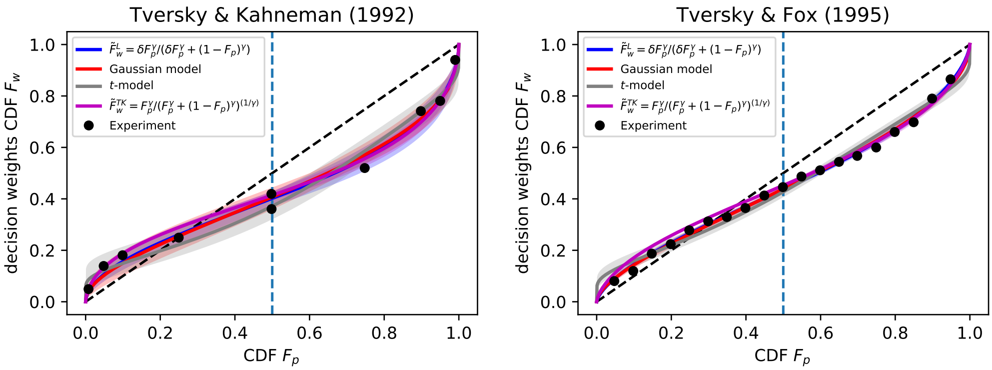
◂ Main Results

Mark Kirstein

Back Up

References

Estimation Error Explains 99% of Probability Weighting



- similar fits of Gaussian & t-distributed model

$\hookrightarrow$ How big is the residual "probability weighting" after accounting for estimation errors?

◂ Main Results

Functional Forms

Gaussian

Tversky and Kahneman (1992, $\gamma = 0.68$)

$$\tilde{F}_w^{TK}\left(F_p;\gamma\right) = \left(F_p\right)^\gamma \frac{1}{\left[\left(F_p\right)^\gamma + \left(1 - F_p\right)^\gamma\right]^{1/\gamma}} \tag{5}$$

Lattimore, Baker, and Witte (1992)

$$\tilde{F}_w^{L}\left(F_p;\delta,\gamma\right) = \frac{\delta F_p^\gamma}{\delta F_p^\gamma + \left(1 - F_p\right)^\gamma} \tag{6}$$

We derive decision weight as a function of probability with $(\alpha\sigma)^2$ as the DM's scale

$$w(p) = p^{\frac{1}{\alpha^2}} \frac{\left(2\pi\sigma^2\right)^{\frac{1-\alpha^2}{2\alpha^2}}}{\alpha}, \tag{7}$$

which is a power law in p with a pre-factor to ensure normalisation

Linking Probability Weighting to Relative Uncertainties

Decision weight w is the normalised sum of the probability $p(x)$ and its uncertainty $\varepsilon\left[p(x)\right]$

$$w(x) = \frac{p(x) + \varepsilon\left[p(x)\right]}{\int_{-\infty}^{\infty} \Big(p(s) + \varepsilon\left[p(s)\right]\Big)\, ds} \,. \tag{8}$$

This can be expressed as

$$w(x) = p(x) \left(\frac{1 + \frac{\varepsilon[p(x)]}{p(x)}}{\int_{-\infty}^{\infty} p(s) \left\{ 1 + \frac{\varepsilon[p(s)]}{p(s)} \right\} ds} \right) \,, \tag{9}$$

where $\frac{\varepsilon[p(x)]}{p(x)}$ is the relative error, which is large (small) for small (large) probabilities

In the long-time limit $w(x) \to p(x)$

References I

Mark Kirstein

Back Up

References

Gigerenzer, Gerd (1991). "How to Make Cognitive Illusions Disappear: Beyond "Heuristics and Biases"". *European Review of Social Psychology* 2 (1), pp. 83–115. DOI:10.1080/14792779143000033 (cit. on p. 12).

Gigerenzer, Gerd (2018). "The Bias Bias in Behavioral Economics". *Review of Behavioral Economics* 5 (3-4), pp. 303–336. DOI:10.1561/105.00000092 (cit. on p. 12).

Hertwig, Ralph and Gerd Gigerenzer (1999). "The 'conjunction fallacy' revisited: how intelligent inferences look like reasoning errors". *Journal of Behavioral Decision Making* 12 (4), pp. 275–305. DOI:10.1002/(sici)1099-0771(199912)12:4¡275::aid-bdm323¿3.0.co;2-m (cit. on p. 12).

Kahneman, Daniel and Amos Tversky (1979). "Prospect Theory: An Analysis of Decision under Risk". *Econometrica* 47 (2), pp. 263–291. DOI:10.2307/1914185 (cit. on p. 21).

Lattimore, Pamela K., Joanna R. Baker, and A. Dryden Witte (1992). "Influence of Probability on Risky Choice: A Parametric Examination". *Journal of Economic Behavior and Organization* 17 (3), pp. 377–400. DOI:10.1016/S0167-2681(95)90015-2 (cit. on p. 23).

Tversky, Amos and Daniel Kahneman (1992). "Advances in Prospect Theory: Cumulative Representation of Uncertainty". *Journal of Risk and Uncertainty* 5 (4), pp. 297–323. DOI:10.1007/BF00122574 (cit. on pp. 3, 23).