What are we weighting for?

A mechanistic model for probability weighting

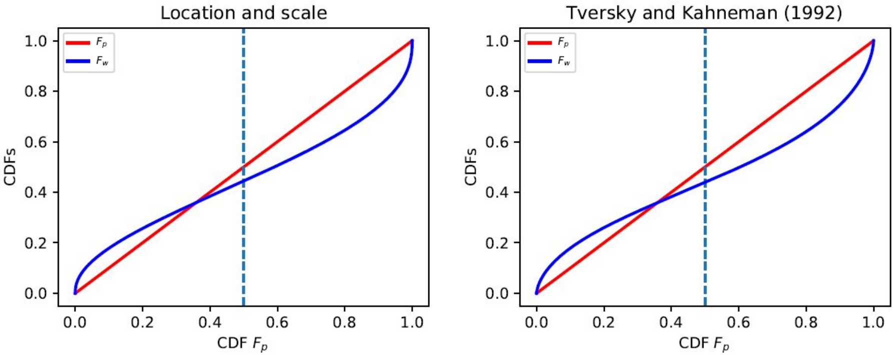
Ole Peters Alexander Adamou Yonatan Berman Mark Kirstein

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

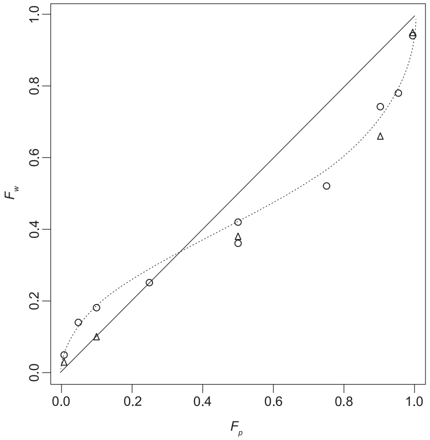


1. generic inverse-S shape can be explained by difference in uncertainty
2. relative estimation error in $p(x)$ is greater for rarer events

▸ PW K&T 1979

Mark Kirstein

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 same random variable X, *e.g.* payout of a gamble

probabilities $p(x)$

CDF $F_p(x)$

Decision Maker (DM)

DM has a model of the same random variable X, *e.g.* payout of a gamble

decision weights $w(x)$

CDF $F_w(x)$

DM's Additional Uncertainty

reasons for DM's extra uncertainty in the random variable X:

- 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,
- ...

DM's Additional Uncertainty

reasons for DM's extra uncertainty in the random variable X:

- 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,
- ...

Let's take a look at the simplest case of extra uncertainty ...

Numerically, our procedure can be applied to arbitrary distributions:

Algo

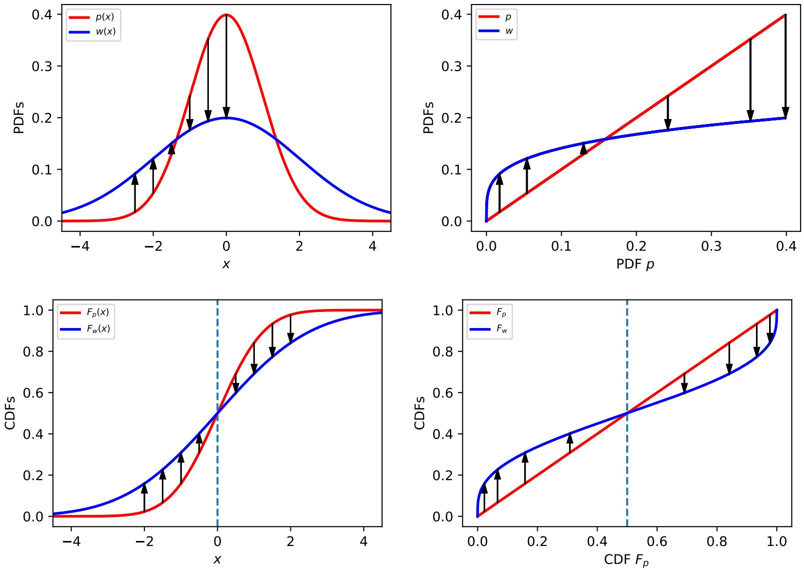
1. construct a list of values for the CDF assumed by the DO, $F_p(x)$
2. construct a list of values for the CDF assumed by the DM, $F_w(x)$
3. plot $F_w(x)$ *vs.* $F_p(x)$

Mark Kirstein

- Main Result
- Probability Weighting
- Uncertainty in PDFs &CDFs
- Ergodicity Question
- Estimation
- Conclusion

Algo

Transmission of different uncertainties from PDFs into CDFs

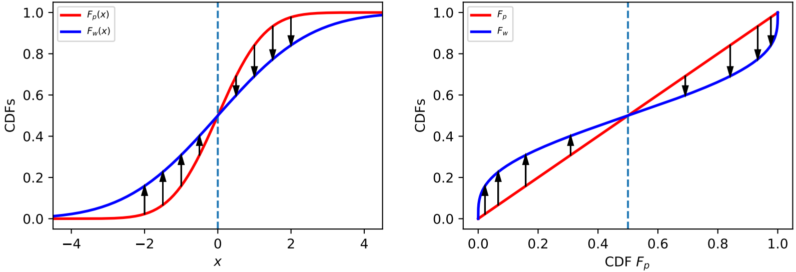


▸ Algo

Transmission of different uncertainties from PDFs into CDFs

Interim conclusion

- greater DM scales reproduces inverse-S shape

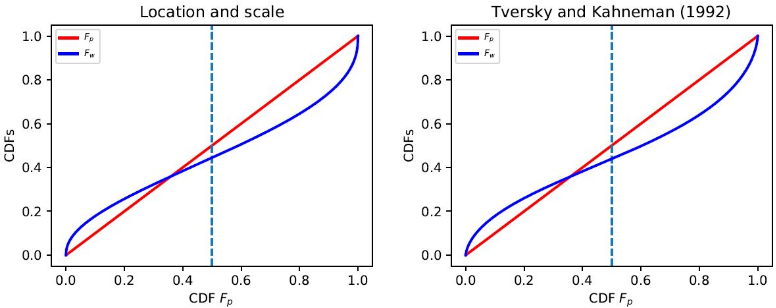


Algo

Transmission of different uncertainties from PDFs into CDFs

Interim conclusion

- greater DM scales reproduces inverse-S shape
- differences in location and scale reproduce asymmetric inverse-S shape
- inverse-S shape whenever DM's assumes greater scale for a unimodal distribution



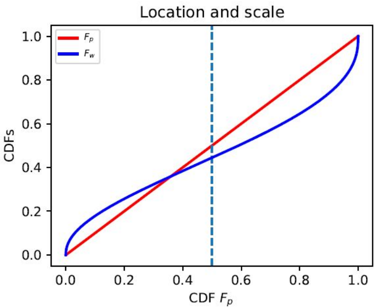
▸ Algo Transmission of different uncertainties from PDFs into CDFs

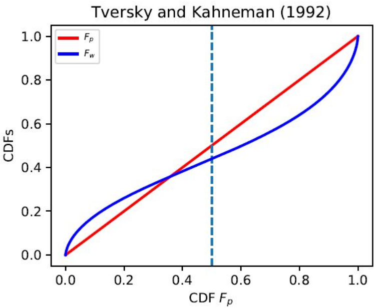
Interim conclusion

- greater DM scales reproduces inverse-S shape
- differences in location and scale reproduce asymmetric inverse-S shape
- inverse-S shape whenever DM's assumes greater scale for a unimodal distribution

- *Job done. Thank you for your attention ;)*

▸ Functional Forms





Asking the Ergodicity Question

Rule: non-ergodic wealth dynamics (*e.g.* multiplicative) (see Peters 2019)

Exception: ergodic wealth dynamics (*e.g.* additive)

DO's concern		**DM's concern**
What happens on average to the ensemble of subjects?	$\neq$	What happens to me on average over time?

- DM's adaptive/ecological rationality $\equiv$ survival, *i.e.* evolutionary incentive to err on the side of caution

$\rightarrow$ add more uncertainty to his model

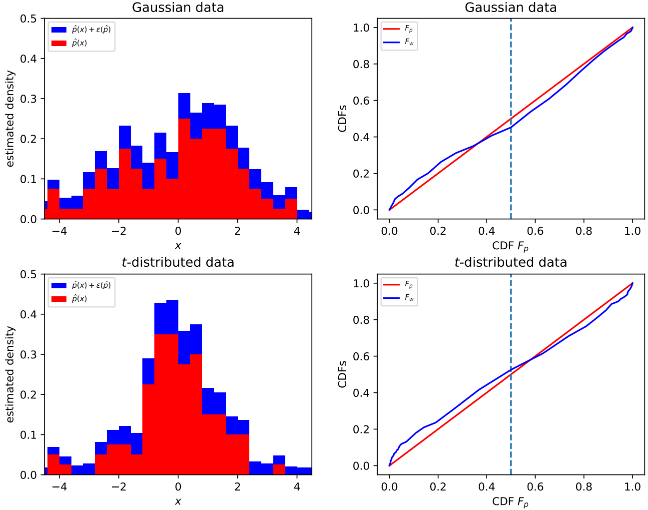
A Mechanism of Estimating p

- usually the DM does not know $p(x)$
- probability is not an observable !!!
- "probability" is polysemous (Gigerenzer 1991, 2018; Hertwig and Gigerenzer 1999)
- DM has to estimate the (unobservable) model parameter $p(x)$ from an observable count of rare events $n(x)$ over time

$\hookrightarrow$ **(Frequentist) Asssumption** $p(x) \equiv$ relative frequency of an event in an infinitely long time series of observation

- very likely to observe no rare events in a finite (small) sample

Simulation of the Estimation



$T = 100$, estimates of $\hat{p}(x)$ in red, estimates with one standard error $\hat{p}(x) + \varepsilon\left[\hat{p}(x)\right]$ in blue

Using the fact that $n(x)$ is a random variable itself, $n(x) \sim \textit{Poisson}$, its fluctuations scale like $\sqrt{n(x)}$
Using the count $n(x)$ to infer the asymptotic PDF as

$$p(x) \approx \frac{n(x)}{T\delta x} \pm \frac{\sqrt{n(x)}}{T\delta x} \tag{1}$$

$$\approx \hat{p}(x) \pm \varepsilon\left[\hat{p}(x)\right] \tag{2}$$

with the standard error (expressed in terms of the estimate itself)

$$\varepsilon\left[\hat{p}(x)\right] \equiv \frac{\sqrt{n(x)}}{T\delta x} = \sqrt{\frac{\hat{p}(x)}{T\delta x}}$$

- standard error $\varepsilon\left[\hat{p}(x)\right]$ shrinks as the probability decreases
- relative error in the estimate is $1/\sqrt{\hat{p}(x)T\delta x}$ grows as the event becomes rarer
- consistent with our claim, that low probabilities come with larger relative errors

$\hookrightarrow$ Errors in probability estimates behave differently for low probabilities than for high probabilities: absolute errors are smaller for lower probabilities, but relative errors are larger

Conclusion

Ergodicity Economics explains probability weighting

- we find an inverse-S shape as a neutral indicator of a difference in opinion
- we find that quite generally the relative uncertainties are larger for rare events than for common events, which generates the inverse-S shape

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

- 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

RESEARCHERS.ONE

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)

↪ distinguish between

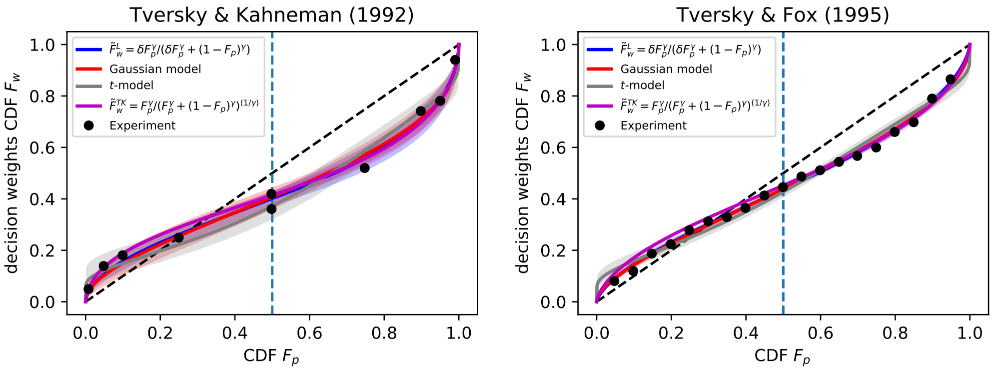
- uncertainty estimation and
- "weighting"

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

↪ How big is the residual "probability weighting" after accounting for uncertainty estimation?

Main Results

Estimation Error Explains 99% of Probability Weighting



- similar fits of our 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, $\gamma = 0.68$)

$$\tilde{F}_w^{TK}\left(F_p;\gamma\right) = \left(F_p\right)^{\gamma} \frac{1}{\left[\left(F_p\right)^{\gamma} + \left(1-F_p\right)^{\gamma}\right]^{1/\gamma}} \tag{3}$$

Lattimore, Baker, and Witte (1992)

$$\tilde{F}_w^{L}\left(F_p;\delta,\gamma\right) = \frac{\delta F_p^{\gamma}}{\delta F_p^{\gamma} + \left(1-F_p\right)^{\gamma}} \tag{4}$$

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

$$w(p) = p^{\frac{1}{\alpha^2}} \frac{\left(2\pi\sigma^2\right)^{\frac{1-\alpha^2}{2\alpha^2}}}{\alpha}\,, \tag{5}$$

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

Different shapes

Numerically, our procedure can be applied to arbitrary distributions:

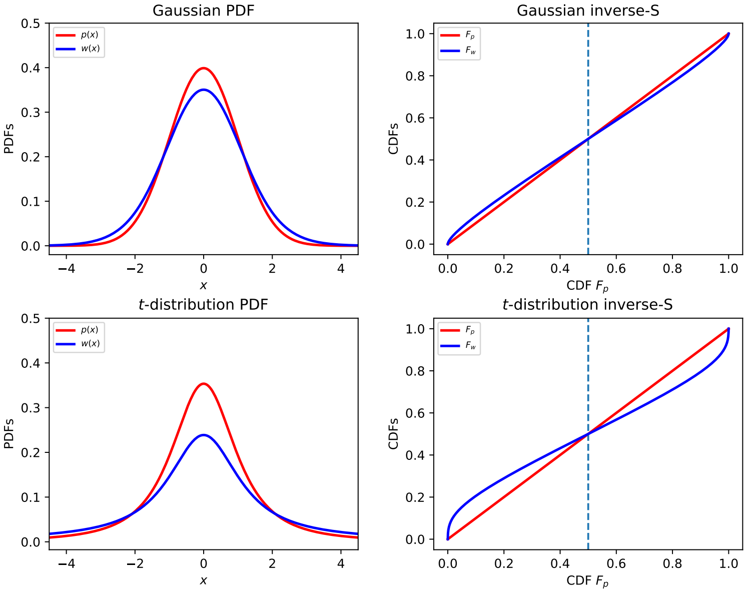
1. construct a list of values for the CDF assumed by the DO, $F_p(x)$
2. construct a list of values for the CDF assumed by the DM, $F_w(x)$
3. plot $F_w(x)$ *vs.* $F_p(x)$

The inverse-S arises whenever a DM assumes a greater scale for a unimodal distribution. To illustrate the generality of the procedure, we carry it out for Student's (power-law tailed) t-distributions (which we refer to as t-distributions), where DO and DM use different shape parameters and different locations

Gaussian Algo

Effect of Different Scales with Heavy-Tailed t-Distributions

Gaussian



Linking Probability Weighting to Relative Uncertainties

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

$$w(x) = \frac{p(x) + \varepsilon\left[p(x)\right]}{\int_{-\infty}^{\infty}\Big(p(s) + \varepsilon\left[p(s)\right]\Big)\,ds}\ . \tag{6}$$

This can be expressed as

$$w(x) = p(x)\left(\frac{1 + \frac{\varepsilon[p(x)]}{p(x)}}{\int_{-\infty}^{\infty} p(s)\left\{1 + \frac{\varepsilon[p(s)]}{p(s)}\right\} ds}\right)\ , \tag{7}$$

where $\frac{\varepsilon[p(x)]}{p(x)}$ is the relative error, which is large (small) for small (large) probabilities
In the long-time limit $w(x) \to p(x)$

Mark Kirstein

Back Up

References

References I

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. 18).

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. 20).

Peters, Ole (2019). "The ergodicity problem in economics". *Nature Physics* 15 (12), pp. 1216–1221. DOI:10.1038/s41567-019-0732-0 (cit. on p. 11).

References II

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, 20).