What are we weighting for?

A mechanistic model for probability weighting

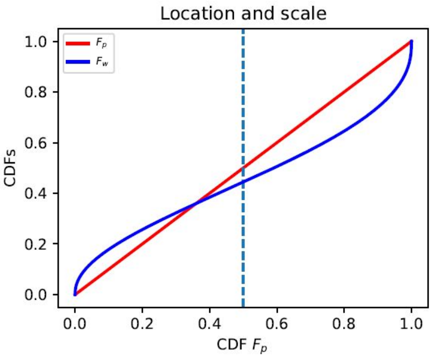
Ole Peters Alexander Adamou Yonatan Berman Mark Kirstein

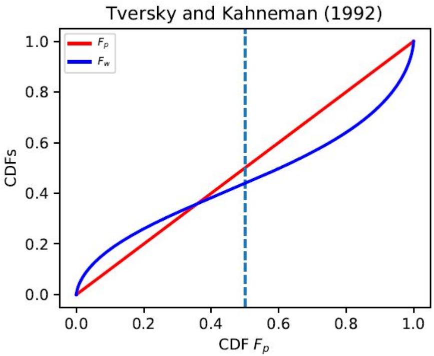
✉ m.kirstein@lml.org.uk
@nonergodicMark

D-TEA 2020, 16 June 2020

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

Main results

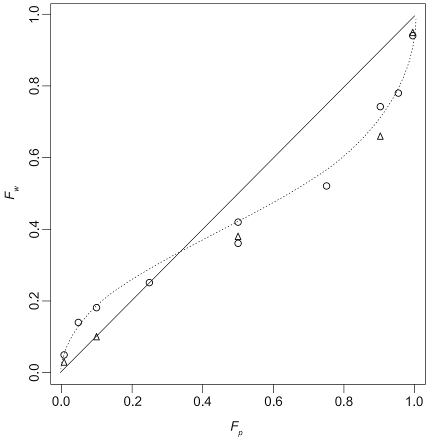




1. generic inverse-S shape can be explained by difference in uncertainty
2. process of estimation of this uncertainty generates inverse-S shape

▸ PW K&T 1979

Probability Weighting (PW)



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

- overestimation of rare events → underestimation of common events
- stable empirical pattern: inverse-S shape

Received wisdom:

- PW = maladaptive irrational cognitive bias

In search of a mechanism

↪ How does this pattern emerge?

↪ 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 $F_p(x)$

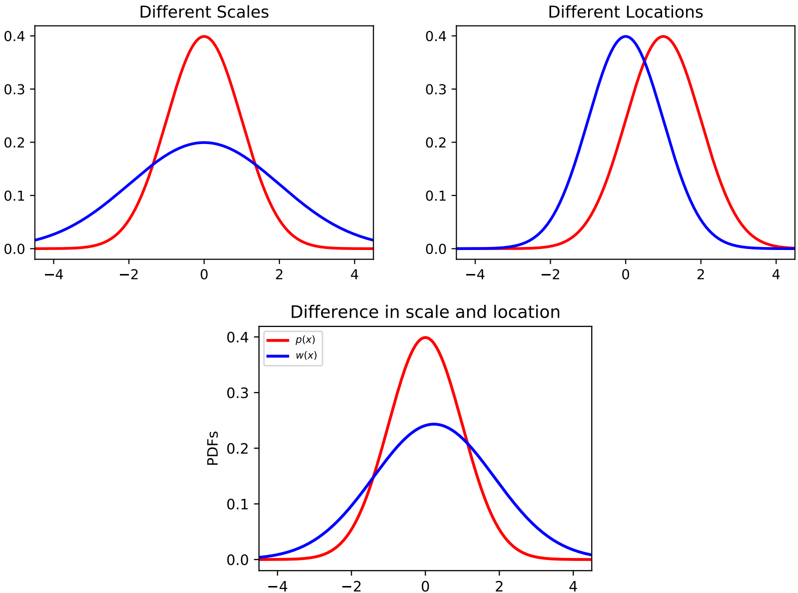
Decision Maker (DM)

DM has a different model of the same random variable X with greater uncertainty

decision weights $w(x)$

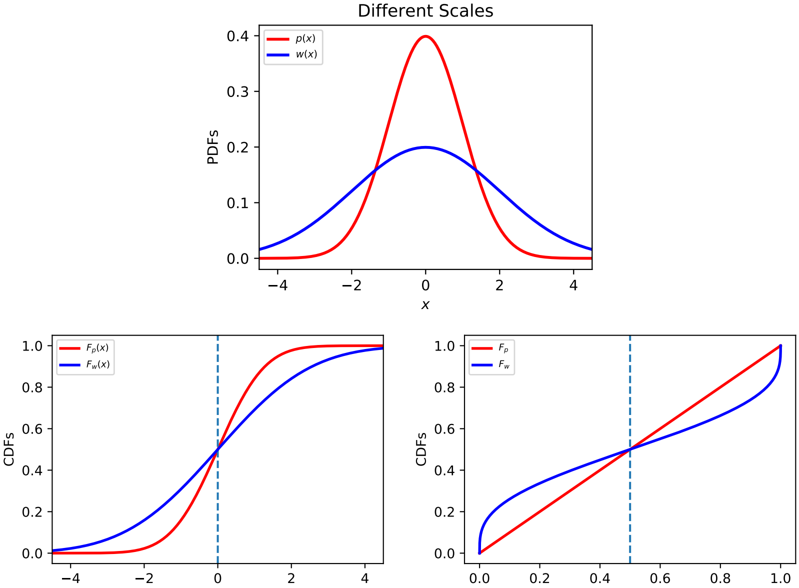
CDF $F_w(x)$

Types of Different Uncertainties

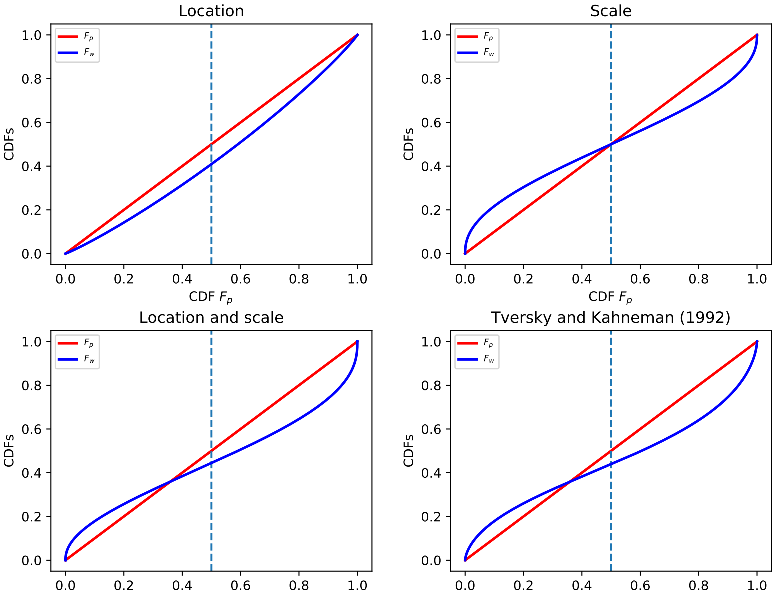


Transmission of Different Uncertainty from PDF in CDF

Algo



Combining Difference in Location and Scale leads to Inverse S



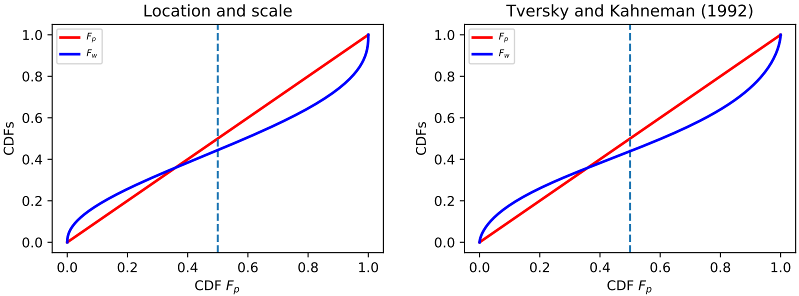
Mark Kirstein

Interim conclusion

- greater scale used by DM 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
- 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?

- 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

Extra Uncertainty is Part of DM's Inference Problem

- "probability" is polysemous (Gigerenzer 1991, 2018; Hertwig and Gigerenzer 1999)
- probabilities are not observable, but
- DM experiences a trajectory of events and
- counts of (rare) events are observable

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

Furthermore ...

- 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,
- ...

Nature of Inference for Rare Events

Rare Event

- asymptotic probability $p(x) = 0.001$
- time series of 100 observations
- $\sim 99.5\%$ of such time series will contain 0 or 1 events
- Naïve estimation: $\hat{p}(x) = 0$ or $\hat{p}(x) = 0.0$, *i.e.* either impossible or ten times more frequently than actual frequency

Common Event

- asymptotic probability $p = 0.5$
- time series of 100 observations
- $\sim 99.5\%$ of time series would contain between 35 and 65 events,
- leading to a much smaller relative error in probability estimates

$\hookrightarrow$ 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

relative estimation errors scales like $^1/_{\sqrt{\text{count}}}$

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: This table assumes $T = 10000$ observed time intervals. To be read as follows (first line): for an event of asymptotic probability 0.1, the most likely count in 10000 trials is 1000. Assuming Poisson statistics, this comes with an estimation error of $\sqrt{1000} = 32$ in the count and $32/10000 = 0.003$ in the probability, which is $0.003/0.1 = 3\%$ of the asymptotic probability.

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

Mark Kirstein

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

RES3ARCHERS.ONE

Mark Kirstein

BACK UP

Mark Kirstein

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

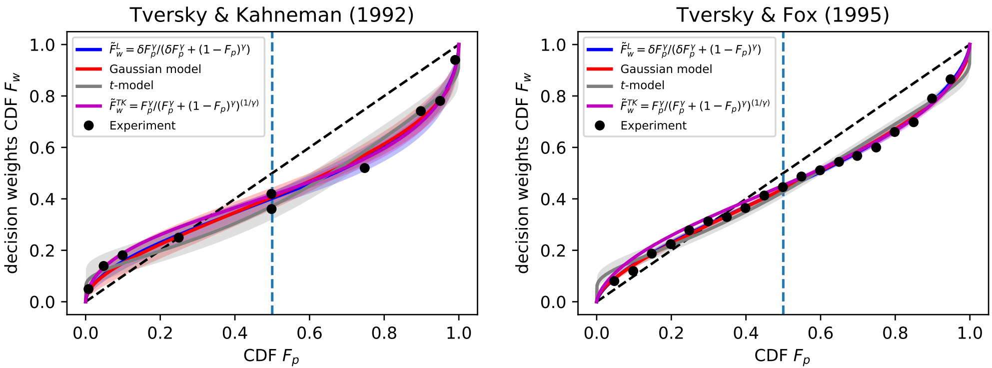
- uncertainty estimation and
- "weighting"

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

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

Mark Kirstein

Back Up

References

Functional Forms

Gaussian

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

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

Lattimore, Baker, and Witte (1992)

$$\tilde{F}_w^{L}\left(F_p;\delta,\gamma\right) = \frac{\delta F_p^{\gamma}}{\delta F_p^{\gamma} + \left(1 - F_p\right)^{\gamma}} \tag{2}$$

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

$$w(p) = p^{\frac{1}{\alpha^2}} \frac{\left(2\pi\sigma^2\right)^{\frac{1-\alpha^2}{2\alpha^2}}}{\alpha}, \tag{3}$$

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

Different shapes

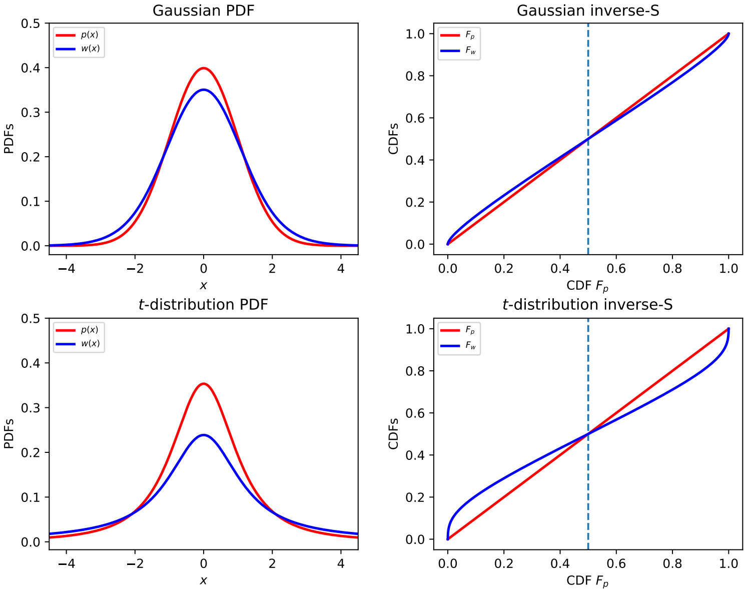
Numerically, our procedure can be applied to arbitrary distributions:

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.* $F_p(x)$

- inverse-S arises for all unimodal distributions
- To illustrate the generality of the procedure, we carry it out for Student's (power-law tailed) t-distributions, where DO and DM use different shape parameters and different locations

Gaussian

Effect of Different Scales with Heavy-Tailed t-Distributions



Linking Probability Weighting to Relative Uncertainties

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

$$w(x) = \frac{p(x) + \varepsilon\left[p(x)\right]}{\int_{-\infty}^{\infty} \Big(p(s) + \varepsilon\left[p(s)\right]\Big)\, ds} \,. \tag{4}$$

This can be expressed as

$$w(x) = p(x) \left(\frac{1 + \frac{\varepsilon[p(x)]}{p(x)}}{\int_{-\infty}^{\infty} p(s) \left\{1 + \frac{\varepsilon[p(s)]}{p(s)}\right\} ds} \right) \,, \tag{5}$$

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. 10).

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. 10).

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. 10).

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. 16).

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. 18).

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, 18).