What are we weighting for?

A mechanistic model for probability weighting

Ole Peters Alexander Adamou Yonatan Berman Mark Kirstein

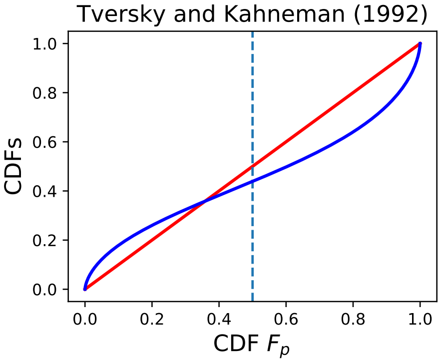
✉ m.kirstein@lml.org.uk
@nonergodicMark

Erich-Schneider-Seminar Economics Department
Christian-Albrechts-Universität zu Kiel
02 November 2020

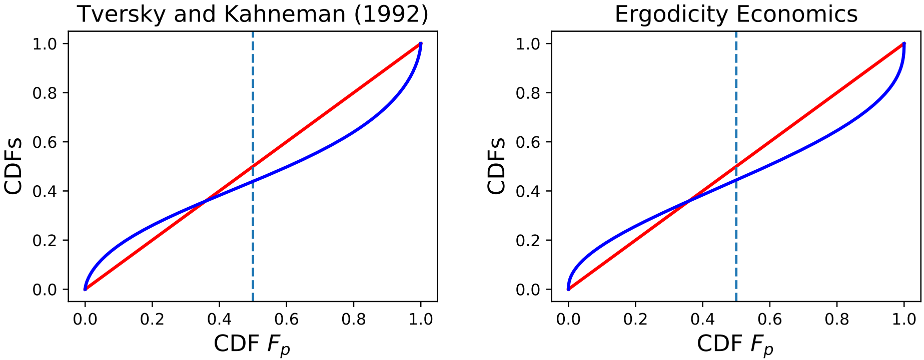
Full paper at https://bit.ly/lml-pw-r1

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

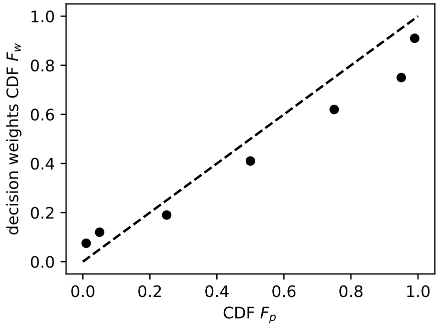
Main results



Main results



Definition of Probability Weighting (PW)



(Preston and Baratta 1948, p. 188, Fig. 1, relabelled axes)

- empirical pattern: inverse-S shape
- Cumulative Prospect Theory (CPT)

Classical interpretation of PW:

- maladaptive irrational cognitive bias

↪ How does this pattern emerge?

↪ Can we derive a functional form (rather than fit a function)?

Related Literature

- TTBOOK no motivation of the functional form, all functions (and parameters therein) used in the literature have no meaning besides producing a fit
- no stable mapping from $p \to w$ (Stewart et al. 2015)
- Ungemach et al. (2009) and Wulff et al. (2018) meta-analysis find all possible weighting curves
- Description-experience gap, experimental design is a key confounder, *i.e.* less overweighting in decisions-from-experience (Hertwig et al. 2004; Hertwig and Erev 2009)

Statistical explanations

-
- PW as a Bayesian heuristic (Martins 2006)

Mark Kirstein

Setup

Task: model payout, x, of a gamble as a random variable.

Disinterested Observer (DO)

DO assigns PDF $p(x)$
$\hookrightarrow$ CDF $F_p(x)$

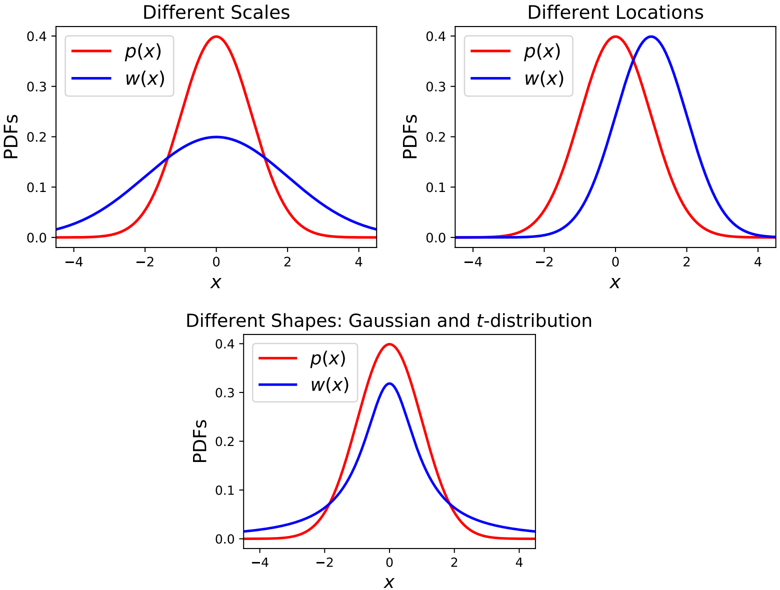
$$F_p(x) = \int_{-\infty}^{x} p(s)\,\mathrm{d}s$$

Decision Maker (DM)

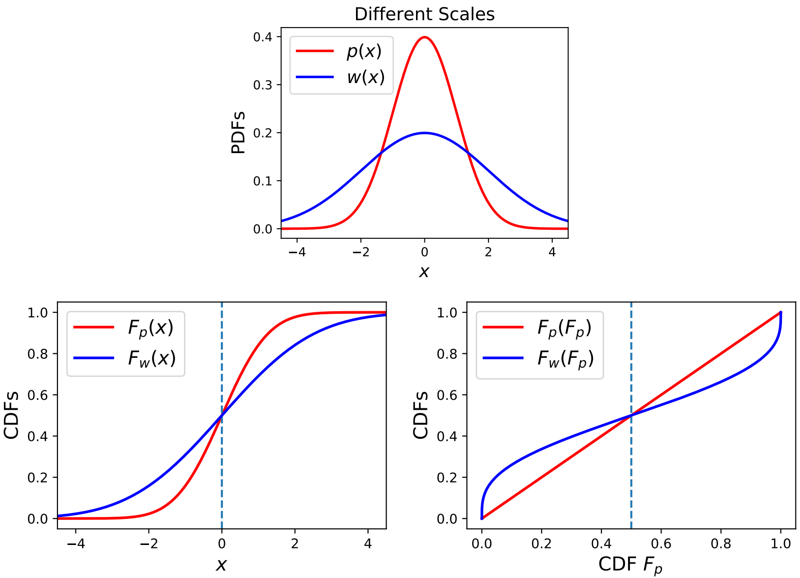
DM assigns different PDF $w(x)$
$\hookrightarrow$ CDF $F_w(x)$

$$F_w(x) = \int_{-\infty}^{x} w(s)\,\mathrm{d}s$$

Scales, Locations, Shapes



Thought Experiment: DM assumes greater scale



Functional form of the weighting function

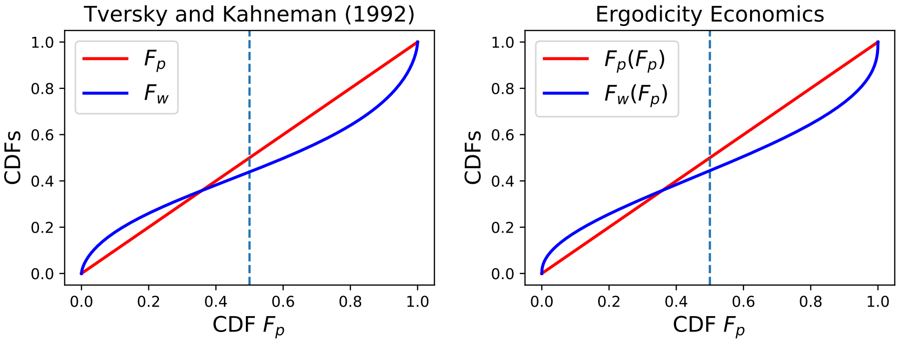
Gaussian case with different scale:

$$w(p) = p^{\frac{1}{\alpha^2}} \underbrace{\frac{\left(2\pi\sigma^2\right)^{\frac{1-\alpha^2}{2\alpha^2}}}{\alpha}}_{\text{normalisation factor}}, \tag{1}$$

where

- DO's scale is σ
- DM's scale is $\alpha\sigma$

Interim conclusion



- DM's greater scale gives inverse-S shape (unimodal distributions)
- difference in locations gives asymmetry
- reproduces observations of probability weighting

Job done. Thank you for your attention ;)

The Ergodicity Question

Typical DO concern

What happens on average to the ensemble of subjects?

$\neq$

Typical DM concern

What happens to me on average over time?

Why DM's greater scale?

- DM has no control over experiment
- experiment may be unclear to DM
- DM may not trust DO
- ...

Experiencing probabilities

- probabilities are not observable
- probabilities encountered as
  - known frequencies in ensemble of experiments (DO)
  - frequencies estimated over time (DM)

$\hookrightarrow$ **estimates have uncertainties – cautious DM accounts for these**

Estimating probabilities

Rare Event

- $p(x) = 0.001$
- 100 observations
- $\sim$ 99.5% get 0 or 1 events
- $\hat{p}(x) = 0$ or $\hat{p}(x) = 0.01$

$\hookrightarrow$ $\hat{p}(x)$ off by 1000%

Common Event

- $p(x) = 0.5$
- 100 observations
- $\sim$ 99.5% get between 35 and 65 events,
- $0.35 < \hat{p}(x) < 0.65$

$\hookrightarrow$ $\hat{p}(x)$ off by 30%

$\hookrightarrow$ small $p(x)$, small count $\rightarrow$ big uncertainty

DMs don't like surprises

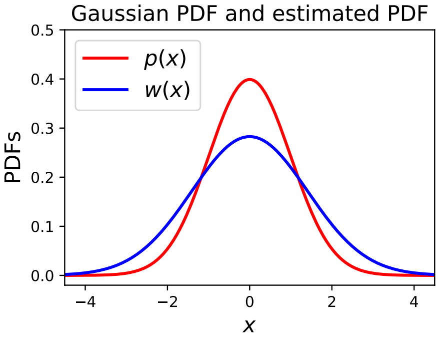
To avoid surprises, DMs add estimation uncertainty $\varepsilon\left[\hat{p}(x)\right]$ to every estimated probability, then normalize, s.t.

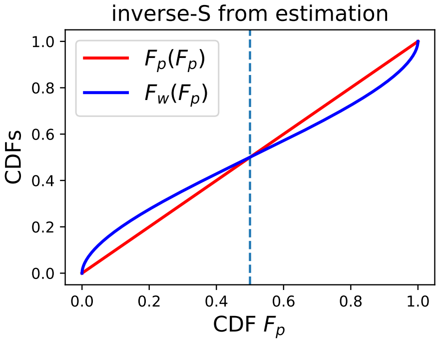
$$w(x) = \frac{\hat{p}(x)+\varepsilon\left[\hat{p}(x)\right]}{\int\left(\hat{p}(s)+\varepsilon\left[\hat{p}(s)\right]\right)ds} \tag{2}$$

DMs don't like surprises

To avoid surprises, DMs add estimation uncertainty $\varepsilon\left[\hat{p}(x)\right]$ to every estimated probability, then normalize, s.t.

$$w(x) = \frac{\hat{p}(x)+\varepsilon\left[\hat{p}(x)\right]}{\int\left(\hat{p}(s)+\varepsilon\left[\hat{p}(s)\right]\right)ds} \quad (2)$$





Conclusion

Classical interpretation of PW

- overestimation of low probability events
- underestimation of high probability events

$\hookrightarrow$ maladaptive irrational cognitive bias

Ergodicity Economics and PW

- inverse-S shape: neutral indicator of different models of the world
- reported observations consistent with DM's extra uncertainty
- may arise from DM estimating probabilities over time

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

- testable prediction $\rightarrow$ Let's run an experiment!
- Manuscript at https://bit.ly/lml-pw-r1
- Interactive code at https://bit.ly/lml-pw-count-b

Conclusion

Classical interpretation of PW

- overestimation of low probability events
- underestimation of high probability events

$\hookrightarrow$ maladaptive irrational cognitive bias

Ergodicity Economics and PW

- inverse-S shape: neutral indicator of different models of the world
- reported observations consistent with DM's extra uncertainty
- may arise from DM estimating probabilities over time

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

- testable prediction $\rightarrow$ Let's run an experiment!
- Manuscript at https://bit.ly/lml-pw-r1
- Interactive code at https://bit.ly/lml-pw-count-b

Thank you for your attention!

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)

- ↪ 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 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 et al. (1992)

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

Gaussian case with greater DM scale $\alpha\sigma$

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

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

References I

Hertwig, Ralph, Greg Barron, Elke U. Weber, and Ido Erev (2004). "Decisions from Experience and the Effect of Rare Events in Risky Choice". *Psychological Science* 15 (8), pp. 534–539. DOI:10.1111/j.0956-7976.2004.00715.x (cit. on p. 5).

Hertwig, Ralph and Ido Erev (2009). "The description–experience gap in risky choice". *Trends in Cognitive Sciences* 13 (12), pp. 517–523. DOI:10.1016/j.tics.2009.09.004 (cit. on p. 5).

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. 20).

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. 22).

Martins, André C. R. (2006). "Probability biases as Bayesian inference". *Judgement and Decision Making* 1 (2), pp. 108–117. URL: `http://journal.sjdm.org/vol1.2.htm` (cit. on p. 5).

Preston, Malcolm G. and Philip Baratta (1948). "An Experimental Study of the Auction-Value of an Uncertain Outcome". *American Journal of Psychology* 61 (2), p. 183. DOI:10.2307/1416964 (cit. on p. 4).

Mark Kirstein

Back Up

References

Stewart, Neil, Stian Reimers, and Adam J. L. Harris (2015). “On the Origin of Utility, Weighting, and Discounting Functions: How They Get Their Shapes and How to Change Their Shapes”. *Management Science* 61 (3), pp. 687–705. DOI:10.1287/mnsc.2013.1853 (cit. on p. 5).

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 p. 22).

Ungemach, Christoph, Nick Chater, and Neil Stewart (2009). “Are Probabilities Overweighted or Underweighted When Rare Outcomes Are Experienced (Rarely)?” *Psychological Science* 20 (4), pp. 473–479. DOI:10.1111/j.1467-9280.2009.02319.x (cit. on p. 5).

Wulff, Dirk U., Max Mergenthaler-Canseco, and Ralph Hertwig (2018). “A meta-analytic review of two modes of learning and the description-experience gap.”. *Psychological Bulletin* 144 (2), pp. 140–176. DOI:10.1037/bul0000115 (cit. on p. 5).