Prospect Theory and the Wisdom of the Inner Crowd

Stephan Hartmann*

July 23, 2017

Abstract

We give a probabilistic justification of the shape of one of the probability weighting functions used in Prospect Theory. To do so, we use an idea recently introduced by Herzog and Hertwig (2014). Along the way we also suggest a new method for the aggregation of probabilities using statistical distances.

Keywords: Prospect Theory, Probability Aggregation, Kullback-Leibler Divergence

1 Introduction

Let p be the objective probability of some event $\mathcal{E}$. According to Prospect Theory people do not use p in their reasoning and decision making, but use a weighted version of p instead. One way of expressing the functional form of the *probability weighting function* is

$$w(p) := \frac{\delta\, p^\gamma}{\delta\, p^\gamma + \overline{p}^\gamma}, \tag{1}$$

with parameter $\gamma, \delta > 0$ (Gonzales 1999) and the short hand $\overline{p} := 1 - p$ which we will use throughout this paper. Fig. 1 illustrates the shape of the curve for various values of the parameters. Prospect theory is descriptively very successful, but it is considered to lack a normative foundation (Wakker 2010).

In this paper we provide a Bayesian justification of this probability weighting function $w(p)$. To do so, we assume that the agent is not content with learning the objective probability of $\mathcal{E}$. After all, it is not clear whether the information source is fully reliable and it is, in any case, a good idea to come up with an independent assessment of the situation. We therefore assume that the agent generates $n-1$ further probability estimates of $\mathcal{E}$, i.e. probabilities $p_1, \ldots, p_{n-1}$ by "harnessing the wisdom of her inner crowd" as suggested in Herzog and Hertwig (2014). She then aggregates these probabilities together with the objective probability p, each with the same weight, to a probability value p'. This is done by minimizing the average Kullback-Leibler divergence between p' and $p, p_1, \ldots, p_{n-1}$. The main result of this paper is that this implies that $p' = w(p)$.

The remainder of this paper is organized as follows. Sec. 2 introduces a new method for the aggregation of probability values. Sec. 3 applies this method to the present case. Finally, Sec. 4 concludes with the suggestion of two lab experiments that can be done to test our proposal.

*Munich Center for Mathematical Philosophy, LMU Munich, 80539 Munich (Germany) – http://www.stephanhartmann.org – s.hartmann@lmu.de.

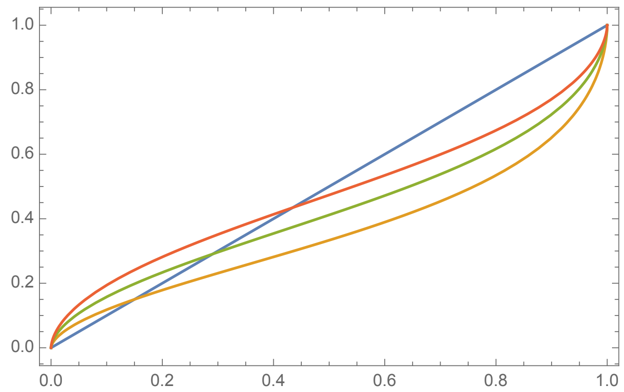

Figure 1: The weighing function $w(p)$ from eq. (6) for $\gamma = .5$ ans $\delta = .5$ (orange), $\delta = .7$ (green) and $\delta = .9$ (red).

2 A New Method for Probability Aggregation

A group of n members has to fix the probability of some event $\mathcal{E}$. After considering the available evidence, each group member i $(i = 1, \ldots, n)$ submits a (subjective) probability value $p_i \in (0, 1)$ to a chairperson who is not part of the group and whose sole task it is to aggregate the probability values of the group members to a group probability value p'. The chairperson considers all group members to be equally reliable and has no additional evidence about the probability value at her disposal. But how should the probabilities be aggregated? What is the rational way to proceed here?

One possibility is to take the *arithmetic mean* of the individual probability values, i.e.

$$\mu_a(p_1, \ldots, p_n) := \frac{1}{n} (p_1 + \cdots + p_n). \tag{2}$$

Another possibility is to take the *geometric mean*, i.e.

$$\mu_g(p_1, \ldots, p_n) := \sqrt[n]{p_1 \cdots p_n}. \tag{3}$$

Note that μ_g leads to a group probability of zero if at least one of the group members submits a probability that approaches zero.

Here we take a different route and determine the value of p' by minimizing the arithmetical average Kullback-Leibler divergence between p' and the p_i's. The Kullback-Leibler divergence between two probability distributions P' and P is defined as follows:

Definition 1 (The Kullback-Leibler Divergence) *Let* $\mathrm{S}_1, \ldots, \mathrm{S}_\mathrm{n}$ *be the possible values of a random variable* S *over which the probability distributions* P *and* P' *are defined. The Kullback-Leibler divergence between* P' *and* P *is then given by*

$$D_{KL}(P'||P) := \sum_{j=1}^{n} P'(\mathrm{S_j}) \log \frac{P'(\mathrm{S_j})}{P(\mathrm{S_j})} .$$

Applied to our case, we want to minimize

$$KL = \frac{1}{n}\sum_{i=1}^{n} D_{KL}(p'||p_i)\,, \tag{4}$$

with

$$D_{KL}(p'||p_i) := p' \log\frac{p'}{p_i} + \overline{p}' \log\frac{\overline{p}'}{\overline{p}_i}. \tag{5}$$

We now define

Definition 2 (The Probabilistic Mean)

$$\mu_p(p_1,\ldots,p_n) := \frac{\mu_g(p_1,\ldots,p_n)}{\mu_g(p_1,\ldots,p_n)+\mu_g(\overline{p}_1,\ldots,\overline{p}_n)}.$$

Then the following theorem holds (all proofs are in the appendix):

Theorem 1 *The probabilistic mean $\mu_p(p_1,\ldots,p_n)$ minimizes KL from eqs. (4) and (5).*

Note that μ_p can be represented more compactly: Let the *individual odds* $o_i := p_i/\overline{p}_i$ and let the *collective odds* $O := \prod_{i=1}^{n} o_i$. Then $\mu_p(p_1,\ldots,p_n) = O/(O+1)$.

Note that $p' := \mu_p(p_1,\ldots,p_n)$ satisfies a number of interesting conditions (the proofs are obvious).

Anonymity $p'(p_1,\ldots,p_n) = p'(\pi(p_1,\ldots,p_n))$ where π is a permutation operator.

All group members contribute equally to the collective probability assignment.

Unanimity Preservation $p'(p,\ldots,p) = p$

For a discussion of this compelling axiom, see Dietrich and List (2014).

Complementarity $p'(p_1,\ldots,p_n) + p'(\overline{p}_1,\ldots,\overline{p}_n) = 1$

If each group member submits the complementary probability value (i.e. $\overline{p} := 1-p$ instead of p), then the group chooses the complementary probability value.

Floor Dominance $\lim_{p_j\to 0} p'(p_1,\ldots,p_n) = 0$ for some $j \in \{1,\ldots,n\}$

If one group member considers the event impossible (and no group member considers the event certain), then the group considers the event impossible. An agent who assigns an event a probability which approaches 0 is prepared to bet any amount of money that this is the right assignment. Otherwise a rational agent would not do it. Hence, it is rational to follow this agent. (Note that all other probabilities are kept fixed and are in $(0,1)$.)

Ceiling Dominance $\lim_{p_j\to 1} p'(p_1,\ldots,p_n) = 1$ for some $j \in \{1,\ldots,n\}$

If one group member considers the event certain (and no group member considers the event impossible), then the group considers the event certain. An agent who assigns an event a probability which approaches 1 is prepared to bet any amount of money that this is the right assignment. Otherwise a rational agent would not do it. Hence, it is rational to follow this agent. (Note that all other probabilities are kept fixed and are in $(0,1)$.)

Note that **Complementarity** and **Floor Dominance** imply **Ceiling Dominance**. Note further that μ_a satisfies **Complementarity** but not **Floor Dominance** and **Ceiling Dominance**. μ_g satisfies **Floor Dominance**, but not **Complementarity** and **Ceiling Dominance**. We will explore these conditions in more detail in future work. We will also extend the proposed method to the aggregation of probability distributions and show, e.g., that it satisfies **External Bayesianity** (Dietrich and List 2016).

3 A Probabilistic Justification of Prospect Theory

With this, we can state the following theorem:

Theorem 2 *$p' = w(p)$ from eq. (1) minimizes the Kullback-Leibler divergence between p' and p, $p_1, \ldots, p_{n-1}$ with $\gamma = 1/n$ and*

$$\delta = \left(\frac{p_1 \cdots p_{n-1}}{\overline{p}_1 \cdots \overline{p}_{n-1}} \right)^{1/n}.$$

This theorem suggests a probabilistic justification of Prospect Theory. To see this, we proceed as follows: (i) The agent learns about the objective probability p of the event $\mathcal{E}$. (ii) She then generates $n-1$ further estimates $p_1, \ldots, p_{n-1}$ of the objective probability, e.g. by "harnessing the wisdom of her inner crowd". (iii) Finally, she aggregates p and $p_1, \ldots, p_{n-1}$, all with the same weight, to a probability p' by minimizing the Kullback-Leibler divergence between p' and $p, p_1, \ldots, p_{n-1}$. As a result, one obtains $p' = w(p)$.
Let us explore $w(p)$ a bit more. The following three propositions hold:

Proposition 1 *$w(p)$ is subadditive, i.e. $w(p) + w(\overline{p}) < 1$, iff $\mu_p(p_1, \ldots, p_{n-1}) < 1/2$.*

Note that experiments suggest that $w(p) + w(\overline{p}) < 1$. A sufficient condition for this inequality to hold is that $p_i < 1/2$ for all $i = 1, \ldots, n-1$, but it is worth noting that all that has to hold is that the probabilistic mean of the probabilities generated by our inner crowd is smaller than 1/2.
Next, we determine the criss-crossing point p_c (see again Fig. 1), which satisfies the condition $p_c = w(p_c)$. We obtain:

Proposition 2 *The criss-crossing point p_c is given by the probabilistic mean $\mu_p(p_1, \ldots, p_{n-1})$.*

Finally, we show that

Proposition 3 *$\delta < 1$ if and only if $p_c < 1/2$.*

That is, $w(p)$ is sub-additive iff the criss-crossing point $p_c < 1/2$.

4 Conclusions

Our main result (Theorem 2) is a first step in the direction of providing a probabilistic justification of Prospect Theory. More work needs to be done (including, e.g. the study of different weights for the objective probability and the generated probabilities and the study of other f-divergencies, see Csiszár (2008)), but we hope to have convinced the reader that it is worth to pursue this an endeavour. Theorem 2 also suggests a number of experimental tests. Here are two questions that can be explored in the lab:

1. It is plausible to assume that participants generate more probabilities if they have more time. So $\gamma = 1/n$ should decrease if one gives participants more time to come up with their probability assessment.

2. If our inner crowd would recommend higher probability values, i.e. if the geometric average of these probabilities would be greater than 1/2, then $p_c > 1/2$. Can we manipulate participants to do so (e.g. by introducing a suitable anchor)?

A Proofs

A.1 Theorem 1

To find the minimum, we differentiate KL by p' and obtain:

$$\frac{\partial KL}{\partial p'} = \log \left[\frac{p'}{\overline{p}'} \cdot \sqrt[n]{\frac{\overline{p}_1 \cdots \overline{p}_n}{p_1 \cdots p_n}} \right] = \log \left[\frac{p'}{\overline{p}'} \cdot \frac{\mu_g(\overline{p}_1, \ldots, \overline{p}_n)}{\mu_g(p_1, \ldots, p_n)} \right].$$

Setting this expression equal to zero, we obtain $p' = \mu_p(p_1, \ldots, p_n)$. ∎

A.2 Theorem 2

Using eqs. (4) and (5), we minimize

$$\begin{aligned} KL &= \frac{1}{n} \left(D_{KL}(P'||P) + \sum_{i=1}^{n-1} D_{KL}(P'||P_i) \right) \\ &= \frac{1}{n} \left((p' \log \frac{p'}{p} + \overline{p}' \log \frac{\overline{p}'}{\overline{p}}) + \sum_{i=1}^{n-1} (p' \log \frac{p'}{p_i} + \overline{p}' \log \frac{\overline{p}'}{\overline{p}_i}) \right). \end{aligned}$$

To do so, we differentiate KL by p' and obtain

$$\frac{\partial KL}{\partial p'} = \log \left[\frac{p'}{\overline{p}'} \cdot \sqrt[n]{\frac{\overline{p}\,\overline{p}_1 \cdots \overline{p}_{n-1}}{p\, p_1 \cdots p_{n-1}}} \right].$$

Setting this expression equal to zero, we obtain

$$\begin{aligned}
p' &= \frac{\sqrt[n]{p\,p_1\cdots p_{n-1}}}{\sqrt[n]{p\,p_1\cdots p_{n-1}}+\sqrt[n]{\overline{p}\,\overline{p}_1\cdots\overline{p}_{n-1}}} \\
&= \frac{\sqrt[n]{p_1\cdots p_{n-1}}\cdot\sqrt[n]{p}}{\sqrt[n]{p_1\cdots p_{n-1}}\cdot\sqrt[n]{p}+\sqrt[n]{\overline{p}_1\cdots\overline{p}_{n-1}}\cdot\sqrt[n]{\overline{p}}} \\
&= \frac{\sqrt[n]{\frac{p_1\cdots p_{n-1}}{\overline{p}_1\cdots\overline{p}_{n-1}}}\cdot\sqrt[n]{p}}{\sqrt[n]{\frac{p_1\cdots p_{n-1}}{\overline{p}_1\cdots\overline{p}_{n-1}}}\cdot\sqrt[n]{p}+\sqrt[n]{\overline{p}}} \\
&= \frac{\left(\frac{p_1\cdots p_{n-1}}{\overline{p}_1\cdots\overline{p}_{n-1}}\right)^{1/n}\cdot p^{1/n}}{\left(\frac{p_1\cdots p_{n-1}}{\overline{p}_1\cdots\overline{p}_{n-1}}\right)^{1/n}\cdot p^{1/n}+\overline{p}^{1/n}} \\
&= \frac{\delta\,p^\gamma}{\delta\,p^\gamma+\overline{p}^\gamma}\,,
\end{aligned}$$

with

$$\gamma := 1/n \quad , \quad \delta := \left(\frac{p_1\cdots p_{n-1}}{\overline{p}_1\cdots\overline{p}_{n-1}}\right)^{1/n} . \quad \blacksquare$$

A.3 Proposition 1

First, we show that $w(p)+w(\overline{p}) < 1$ iff $\delta < 1$. This follows from the fact that the following inequalities are equivalent for $p \in (0,1)$ and $\gamma, \delta > 0$:

$$\begin{aligned}
w(p)+w(\overline{p}) &< 1 \\
\frac{\delta\,p^\gamma}{\delta\,p^\gamma+\overline{p}^\gamma}+\frac{\delta\,\overline{p}^\gamma}{\delta\,\overline{p}^\gamma+p^\gamma} &< 1 \\
\delta\,p^\gamma\,(\delta\,\overline{p}^\gamma+p^\gamma)+\delta\,\overline{p}^\gamma\,(\delta\,p^\gamma+\overline{p}^\gamma) &< (\delta\,\overline{p}^\gamma+p^\gamma)(\delta\,p^\gamma+\overline{p}^\gamma) \\
\delta^2 &< 1 \\
\delta &< 1
\end{aligned}$$

Next, we use the definition of δ and show that the following inequalities are equivalent for $p \in (0,1)$:

$$\begin{aligned}
\delta &< 1 \\
\left(\frac{p_1\cdots p_{n-1}}{\overline{p}_1\cdots\overline{p}_{n-1}}\right)^{1/n} &< 1 \\
\left(\frac{p_1\cdots p_{n-1}}{\overline{p}_1\cdots\overline{p}_{n-1}}\right)^{1/(n-1)} &< 1 \\
\frac{\mu_g(\overline{p}_1,\ldots,\overline{p}_{n-1})}{\mu_g(p_1\cdots p_{n-1})} &> 1
\end{aligned}$$

Finally, we note that

$$\begin{aligned}
\mu_p(p_1,\ldots,p_{n-1}) &= \left(1+\frac{\mu_g(\overline{p}_1,\ldots,\overline{p}_{n-1})}{\mu_g(p_1\cdots p_{n-1})}\right)^{-1} \\
&< 1/2. \quad \blacksquare
\end{aligned}$$

A.4 Proposition 2

The following equations are equivalent:

$$\begin{aligned} p_c &= \frac{\delta\, p_c^{\gamma}}{\delta\, p_c^{\gamma} + \overline{p}_c^{\gamma}} \\ \delta\, p_c^{\gamma}\, \overline{p}_c &= p_c\, \overline{p}_c^{\gamma} \\ \delta\, \overline{p}_c^{\overline{\gamma}} &= p_c^{\overline{\gamma}} \\ \delta &= \left(\frac{p_c}{\overline{p}_c}\right)^{\overline{\gamma}}. \end{aligned}$$

From this we conclude that

$$\begin{aligned} p_c &= \frac{\delta^{1/\overline{\gamma}}}{1+\delta^{1/\overline{\gamma}}} \\ &= \frac{\delta^{n/(n-1)}}{1+\delta^{n/(n-1)}} \\ &= \frac{\sqrt[n-1]{p_1\cdots p_{n-1}}}{\sqrt[n-1]{p_1\cdots p_{n-1}} + \sqrt[n-1]{\overline{p}_1\cdots \overline{p}_{n-1}}} \\ &= \frac{\mu_g(p_1,\ldots,p_{n-1})}{\mu_g(p_1,\ldots,p_{n-1}) + \mu_g(\overline{p}_1,\ldots,\overline{p}_{n-1})} \\ &= \mu_p(p_1,\ldots,p_{n-1}). \quad \blacksquare \end{aligned} \tag{6}$$

A.5 Proposition 3

From eq. (6) of the proof of Proposition 2, we obtain that the following inequalities are equivalent:

$$\begin{aligned} p_c &< 1/2 \\ \frac{\delta^{n/(n-1)}}{1+\delta^{n/(n-1)}} &< 1/2 \\ \delta^{n/(n-1)} &< 1/2 + 1/2\,\delta^{n/(n-1)} \\ \delta^{n/(n-1)} &< 1 \\ \delta &< 1 \quad \blacksquare \end{aligned}$$

References

Csiszár (2008). Axiomatic Characterizations of Information Measures. *Entropy* 10: 261–273.

Dietrich, F. and C. List (2016). Probabilistic Opinion Pooling. In: A. Hájek and C. Hitchcock (eds.): *The Oxford Handbook of Probability and Philosophy.* New York: Oxford University Press, ch. 25.

Gonzales, R. (1999). On the Shape of the Probability Weighting Function. *Cognitive Psychology* 38: 129–166.

Herzog, S. and R. Hertwig (2014). Harnessing the Wisdom of the Inner Crowd. *Trends in Cognitive Sciences* 18(10): 504–506.

Wakker, P. (2010). *Prospect Theory.* Cambridge: Cambridge University Press.