

# “Weight of Evidence, Evidential Completeness and Accuracy”

Rafal Urbaniak and Marcello Di Bello

## 1 Motivations

### 1.1 Balance vs. weight

Suppose we want to represent our uncertainty about a proposition in terms of a single probability that we assign to it. It is not too difficult to inspire the intuition that this representation does not capture an important dimension of how our uncertainty connects with the evidence we have or have not obtained. In a 1872 manuscript of *The Fixation of Belief* (W3 295) C. S. Peirce gives an example meant to do exactly that.

When we have drawn a thousand times, if about half have been white, we have great confidence in this result. We now feel pretty sure that, if we were to make a large number of bets upon the color of single beans drawn from the bag, we could approximately insure ourselves in the long run by betting each time upon the white, a confidence which would be entirely wanting if, instead of sampling the bag by 1000 drawings, we had done so by only two.

The objection is not too complicated. Your best estimate of the probability of  $W$  = ‘the next bean will be white’ is .5 if half of the beans you have drawn randomly so far have been white, no matter whether you have drawn a thousand or only two of them. But this means that expressing your uncertainty about  $W$  by locutions such as ‘my confidence in  $W$  is .5’ does not capture this intuitively important distinction.

Similar remarks can be found in Peirce’s 1878 *Probability of Induction*. There, he also proposes to represent uncertainty by at least two numbers, the first depending on the inferred probability, and the second measuring the amount of knowledge obtained; as the latter, Peirce proposed to use some dispersion-related measure of error (but then suggested that an error of that estimate should also be estimated and so, so that ideally more numbers representing errors would be needed).

Peirce himself did not call this the weight of evidence (and in fact, used the phrase rather to refer to the balance of evidence, W3 294) [CITE KASSER 2015]. However, his criticism of such an oversimplified representation of uncertainty anticipated what came to be called weight of evidence by Keynes in his 1921 *A Treatise on Probability*:

As the relevant evidence at our disposal increases, the magnitude of the probability of the argument may either increase or decrease, according as the new knowledge strengthens the unfavourable or the favourable evidence; but something seems to have increased in either case,—we have a more substantial basis upon which to rest our conclusion. I express this by saying that an accession of new evidence increases the weight of an argument. New evidence will sometimes decrease the probability of an argument but it will always increase its ‘weight.’ (p. 71)

The key point is the same [CITE LEVI 2001]: the balance of probability alone cannot characterize all important aspects of evidential appraisal. Keynes also considered measuring weight of evidence in terms of the variance of the posterior distribution of a certain parameter, but was quite attached to the idea that weight should increase with new information, even if the dispersion increase with new evidence [TP 80-82], and so he proposed only a very rough sketch of a positive sketch. Moreover, as he was uncertain how a measure of weight should be incorporated in further decision-making, he was skeptical about the practical significance of the notion. [TP 83]

But what is this positive sketch? On one hand, Keynes [TP 58-59] connects the notion of weight with relevance. Call evidence  $E$  relevant to  $X$  given  $K$  just in case  $\Pr(X|K \wedge E) \neq \Pr(X|K)$ .<sup>1</sup> One postulate than can be found in the *Treatise* [TP 84] is:<sup>2</sup>

(R-monotonicity) If  $E$  is relevant to  $X$  given  $K$ , where  $K$  is background knowledge,  $V(X|K \wedge E) > V(X|K)$ , where  $V$  is the weight of evidence.

[RUNDE 1990, 280] suggests that Keynes at some point calls weight the completeness of information. This however, is a bit hasty, as Keynes only says that *the degree of completeness of the information on which a probability is based does seem to be relevant, as well as the actual magnitude of the probability, in making practical decisions*. As later on we will argue that it is actually useful to distinguish evidential weight (how much evidence do we have?) and evidential completeness (do we have all the evidence that we would expect in a given case?), we rather prefer to extract a more modest postulate:

(Completeness) If  $E_1$  and  $E_2$  are relevant items of evidence, and  $E_2$  is (in a sense to be discussed) more complete than  $E_1$ ,  $V(X|K \wedge E_2) > V(X|K \wedge E_1)$ .

If we conceptualize  $E_2$  being complete and  $E_1$  being incomplete as  $E_2$  being a maximal relevant conjunction of relevant claims one of which is  $E_1$ , (Completeness) follows from (R-monotonicity).

Now, some requirements on how weight of evidence is related to the balance of probability. For one thing, Keynes insists that new (relevant) evidence might decrease probability but will always increase weight [TP 77]. Since (R-monotonicity) already captures the idea that weight will always increase, here we extract the other part of the claim:

(Possible decrease) It is possible that  $V(X|K \wedge E) > V(X|K)$  while  $P(X|K \wedge E) < P(X|K)$ .

Clearly, Keynes also endorsed the following two requirements of a very similar form:

(Possible increase) It is possible that  $V(X|K \wedge E) > V(X|K)$  while  $P(X|K \wedge E) > P(X|K)$ .

(Possibly no change) It is possible that  $V(X|K \wedge E) > V(X|K)$  while  $P(X|K \wedge E) = P(X|K)$ .

Keynes is not referring to the sheer number of statements on the right hand side of a conditional probability  $P(H|E)$  or the sheer bulk of information that these statements contain. By “relevant evidence,” Keynes is only referring to the extent that  $E$  provides information that is pertinent to  $H$  in particular. [Pedden3]

## 1.2 Examples and informal desiderata

- Go over Nance in particular, Cohen, some other sources?

### 1.2.1 Monotonicity of weight

Runde, Joyce, Weatherson, Peden

## 1.3 Hamer’s weight of evidence

## 1.4 Good’s weigh of evidence and the information value

- present Good, discuss
- compare to pointwise mutual information
- evaluate in light of the desiderata

## 1.5 Skyrms and resilience?

## 1.6 Imprecision and weight with intervals

Keynes’ later works and Peden’s paper

<sup>1</sup>As discussed in [CITE RUNDE 1990], Keynes also uses a slightly more convoluted notion of relevance to avoid equally strong items of opposite evidence turning out to be irrelevant; here we do not need to be concerned with this complication.

<sup>2</sup>RUNDE 1990 283 suggests Keynes allows for weight of evidence to decrease when new evidence increases the range of alternatives, but this is based on Keynes’ claim that weight is increased when the number of alternatives is reduced, and Keynes does not directly say anything about the possibility of an increase of the number of alternatives.

**1.6.1 Sharpening by richness**

**1.6.2 Sharpening by specificity**

**1.6.3 Sharpening by precision**

**1.7 Imprecision: a second-order approach**

**1.8 Information-theoretic weight of evidence**

**1.9 Completeness tends to improve weight**

**1.10 Weight tends to improve accuracy**

## **2 Literature to discuss**

Kasser, 2016, Two Conceptions of Weight of Evidence in Peirce's Illustrations of the Logic of Science [DOWNLOADED]

Feduzi, 2010, On Keynes's conception of the weight of evidence [READ]

Cohen 1986, Twelve Questions about Keynes's Concept of Weight [READ]

Pedden, William 2018, Imprecise probability and the measurement of Keynes' weight of arguments

Levi 2011, the weight of argument [DOWNLOADED]

Skyrms 1977 resiliency, propensities [DOWNLOADED]

Synthese 186 (2) 2012, volume on Keynesian weight [CHECKED, NOT MUCH ON WEIGHT ACTUALLY, NO NEED TO READ]

Good, weight of evidence, survey

Good, PROBABILITY AND THE WEIGHING OF EVIDENCE

David Hamer, Probability, anti-resilience, and the weight of expectation [READ]

William Peden, Imprecise Probability and the Measurement of Keynes's "Weight of Arguments"

Runde, Keynesian Uncertainty and the weight of arguments [DOWNLOADED]

Weatherson, 2002, Keynes, uncertainty and interest rates [DOWNLOADED]

Jeffrey M. Keisler, Value of information analysis: the state of application

Edward C. F. Wilson, A Practical Guide to Value of Information Analysis

Joyce JM (2005) How probabilities reflect evidence.