Bayes By the Sea

Formal Epistemology, Statistics and Probability

Ancona, 13th - 14th September 2018
Università Politecnica delle Marche – Facoltà di Economia "Giorgio Fuà"

Roberto Festa

Department of Humanistic Studies
University of Trieste

LAWS OF LIKELIHOOD AND MATTHEW EFFECTS FOR BAYESIAN CONFIRMATION

Table of contents

The starting point of the present research

Festa, Roberto 2012. '"For Unto Every One That Hath Shall Be Given": Matthew Properties for Incremental Confirmation', *Synthese* **184**, pp. 89–100.

Festa, Roberto & Gustavo Cevolani 2017. 'Unfolding the Grammar of Bayesian Confirmation: Likelihood and Antilikelihood Principles', *Philosophy of Science* **84**, no. 1, pp. 56–81.

1. Likelihood principles for Bayesian confirmation

We will focus on hypotheses H and pieces of evidence E with non-extreme probability values, that is, such that

$$0 < p(H), p(E) < 1.$$

The only exception will be the tautological evidence $\top$.

Incremental measures of confirmation are characterized by the following basic principles P1-P3 (Crupi 2015):

P1 ***Formality***

There is a function f such that, for any H and E, $C(H,E) = f[p(H|E), p(H), p(E)]$.

P2 ***Tautological evidence***

For any H1 and H2, $C(H1, \top) = C(H2, \top)$.

P3 ***Final probability***

For any H, E1, E2, if $p(H|E1)$ =/> $p(H|E2)$, then $C(H,E1)$ =/> $C(H,E2)$.

P1 states that C(H,E) only depends on the probability distribution over the algebra generated by H and E.

Such a distribution can be determined by different triples of probability values, [p(H|E), p(H), p(E)] being only one of them.

For instance, P1 can be reformulated, so to speak, in a *likelihood form* as follows:

L-P1 There is a function k such that, for any H and E, C(H,E) = k[p(E), p(E|H), p(E|¬H)].

L-P1 says that C(H,E) can be expressed in *evidential terms*, that is, as a function of the initial probability p(E) of evidence E and of its conditional probabilities p(E|H) and p(E|¬H).

The label "L-P1" is justified by noting that p(E) can be construed as the likelihood p(E|⊤) of a tautological hypothesis.

Accordingly, L-P1 says that C(H, E) depends only on the three likelihoods p(E|H), p(E|¬H), and p(E).

Since p(E|¬H) is *the likelihood of the negation* of the relevant hypothesis H, we will refer to it as the *negation likelihood* – for short, the ***N-likelihood*** –, of H.

The likelihoods p(E|H), p(E|¬H), and p(E) (= p(E|⊤)) can be seen as the measures of the *predictive power* w.r.t. E of the hypothesis H, its negation ¬H and a tautological hypothesis ⊤, respectively.

Incremental measures satisfy the basic principles P1–P3 and all the other principles that can be logically derived from them. Such principles isolate the properties characterizing all incremental measures and are thus labeled *universal principles*.

Certain (classes of) incremental measures may exhibit further interesting properties, which are specified by corresponding *structural principles*.

An important feature of the ***grammar of confirmation*** is the distinction between *universal properties* of confirmation (characterizing all incremental measures) and *structural properties* (isolating specific classes of such measures).

A likelihood principle – for short, ***L-principle*** –, requires certain relationships to hold between C(H,E) and the likelihoods of H and ¬H with respect to E, that is, between C(H,E) and the probabilities p(E), p(E|H) and p(E|¬H).

An *universal L-principle* is the above principle L-P1, saying that there exists a function f such that, for any H and E, C(H,E) = k[p(E), p(E|H), p(E|¬H)].

One may state certain *structural L-principles* which strengthen L-P1 by specifying *how* the value of C(H,E) = k[p(E), p(E|H), p(E|¬H)] changes when two of the likelihoods p(E), p(E|H), and p(E|¬H) are fixed and the other changes.

For instance, we may ask how C(H,E) depends on p(E|H), i.e., how C(H,E) changes when p(E) and p(E|¬H) are fixed and p(E|H) changes. A plausible answer to this question is expressed by the following L-principle:

L-Inc *Confirmation increases with likelihood*
For fixed values of p(E) and p(E|¬H), when p(E|H) increases C(H,E) increases.

Since p(E|H) can be seen as a measure of the *predictive power* of H w.r.t. E, L-Inc amounts to saying that C(H,E) increases together with the predictive power of H.

Below we will focus on the following question:

how does $C(H,E)$ depend on the N-likelihood $p(E|\neg H)$?

In other words,

how does $C(H,E)$ change when $p(E)$ and $p(E|H)$ are fixed and the N-likelihood $p(E|\neg H)$ changes?

Three different answers to this question are provided by the following *N-likelihood principles* – for short, ***N-principles*** (Festa & Cevolani 2017):

N-Ind *Confirmation does not depend on N-likelihood*
For fixed values of p(E) and p(E|H), C(H,E) does not depend on p(E|¬H).

N-Dec *Confirmation decreases with N-likelihood*
For fixed values of p(E) and p(E|H), when p(E|¬H) increases C(H,E) decreases.

N-Inc *Confirmation increases with N-likelihood*
For fixed values of p(E) and p(E|H), when p(E|¬H) increases C(H,E) increases.

The intuitive content of the above N-principles can be expressed as follows.

N-Ind says that C(H,E) *is not affected* by the predictive power p(E|¬H) of the negation of H.

N-Dec says that C(H,E) *is a decreasing function* of p(E|¬H).

N-Inc says that C(H,E) *is an increasing function* of p(E|¬H).

At first sight N-Inc, i.e., the request that C(H,E) increases with the predictive power of the negation of H, seems quite implausible.

However, it has been shown that a number of intuitively well-grounded incremental measures satisfy N-Inc (Festa & Cevolani 2017).

2. Matthew effects for Bayesian confirmation

The basic principle P1 can be reformulated in the following *initial probability form*:

I-P1 There is a function g such that, for any H and E, C(H,E) = g[p(H), p(E), p(E|H)].

I-P1 says that C(H,E) can be expressed as a function of the initial probability p(H) of H and the likelihoods p(E) and p(E|H).

An *initial probability principle* – for short, ***I-principle*** –, requires certain relationships to hold between C(H,E), the initial probability p(H), and the likelihoods p(E) and p(E|H).

While the universal principle I-P1 only says that there is a function g such that, for any H and E, C(H,E) = g[p(H), p(E), p(E|H)], we can state certain ***structural I-principles*** specifying ***how*** the value of C(H,E) changes when two the probabilities p(H), p(E), and p(E|H) are fixed and the other changes.

For our discussion of I-principles, the notions of (*predictive*) *success*, (*predictive*) *failure*, and (*confirmational*) *impact* of successes and failures will be useful.

Predictive success

E is a (predictive) ***success*** of H ≡ $p(E|H) > p(E)$ (equivalently, $p(H|E) > p(H)$, i.e., E confirms H).

Predictive failure

E is a (predictive) ***failure*** of H ≡ $p(E|H) < p(E)$ (equivalently, $p(H|E)$ z $p(H)$, i.e., E disconfirms H).

Confirmational impact of a success

The (confirmational) ***impact of a success*** E on H increases when C(H,E) increases.

Confirmational impact of a failure

The (confirmational) ***impact of a failure*** E on H decreases when C(H,E) increases.

The last definition is suggested by the circumstance that the ***amount of disconfirmation*** that H receives from a piece of evidence E may be identified with some decreasing function of C(H,E).

One can state a number I-principles applying to successes (***IS-principles***) or to failures (***IF-principles***). First of all, let us consider the following three IS- principles:

IS-Ind *Confirmation by a success does not depend on initial probability*
For fixed values of p(E) and p(E|H), with p(E|H) > p(E), when p(H) increases C(H,E) does not change.

IS-Inc *Confirmation by a success increases with initial probability*
For fixed values of p(E) and p(E|H), with p(E|H) > p(E), when p(H) increases C(H,E) increases.

IS-Dec *Confirmation by a success decreases with initial probability*

For fixed values of p(E) and p(E|H), with p(E|H) > p(E), when p(H) increases C(H,E) decreases.

The above IS-principles are logically equivalent to – or closely related with –, the so-called ***Matthew effects***, recently investigated in the literature on Bayesian confirmation (see Kuipers 2000, Festa 2012, Roche 2014).

Now let us consider the following IF-principles corresponding to the above IS-principles:

IF-Ind *Confirmation by failure does not depend on initial probability*
For fixed values of p(E) and p(E|H), with p(E|H) < p(E), when p(H) increases C(H,E) does not change.

IF-Inc *Confirmation by failure increases with initial probability*
For fixed values of p(E) and p(E|H), with p(E|H) < p(E), when p(H) increases C(H,E) increases.

IF-Dec *Confirmation by failure decreases with initial probability*
For fixed values of p(E) and p(E|H), with p(E|H) < p(E), when p(H) increases C(H,E) decreases.

Some interesting logical relations hold among the six I-principles above:

- The three IS-principles above are *incompatibles* with each other and the same holds for the three IF-principles above.

- Each of the three IS-principles above is ***compatible*** with each of the three IF-principles above.

In spite of the above mentioned logical compatibility, it should be noted that some pairs formed by an IS-principle and a corresponding IF-principle are *intuitively more natural* than other pairs.

In particular, let us consider the following three *impact principles* – for short, ***Imp-principles***:

Imp-Ind	IS-Ind & IF-Ind
Imp-Inc	IS-Inc & IF-Dec
Imp-Dec	IS- Dec & IF-Inc

The term "impact principles" for the three Imp-principles above is suggested by the circumstance that each of them is a conjunction formed by an IS-principle and an IF-principle expressing identical intuitions about the impact of successes and failures.

Indeed, the above Imp-principles can be informally restated in terms of impact as follows:

Imp-Ind *Impact does not depend on initial probability*
The impact of E on H is not affected by p(H).

Imp-Inc *Impact increases with initial probability*
The impact of E on H increases when p(H) increases.

Imp-Dec *Impact decreases with initial probability*
The impact of E on H decreases when p(H) increases.

3. Equivalence between likelihood principles and Matthew effects

Above we have considered two kinds of principles for incremental confirmation, i.e., the likelihood principles, or L-principles, and the initial probability principles, or I-principles.

Recall that

- a **L-principle** requires certain relationships to hold between C(H,E) and the three likelihoods p(E), p(E|H) and p(E|¬H)

- an **I-principle** principles requires certain relationships to hold between C(H,E), the initial probability p(H), and the two likelihoods p(E), and p(E|H).

This means that

- both L- and I-principles require that certain relationships hold between C(H,E) and the two likelihoods p(E), and p(E|H).

- a ***crucial difference*** between L- and I-principles is given by the circumstance that the former specify how C(H,E) depends on the N-likelihood p(E|¬H) while the latter specify how C(H,E) depends on the initial probability p(H).

Possibly due to the above crucial difference, the possibility that some L-principles are equivalent to some I-principles has been neglected.

However, this possibility should be carefully considered. In particular, one should consider

- the possibility that there is some equivalence between a certain kind of L-principles, i.e., the N-likelihood principles, and a certain kind of I-principles, i.e., the Imp-principles.

Recall that

- the N-principles specify how C(H,E) changes when the two likelihood p(E) and p(E|H) are fixed and *the N-likelihood p(E|¬H) changes*;

- the Imp-principles specify how C(H,E) changes when the two likelihood p(E) and p(E|H) are fixed and *the initial probability p(H) changes*;

- hence, the possibility of some equivalence between N-principles and Imp-principles depends on the possibility that, for fixed values of p(E) and p(E|H), *the N-likelihood p(E/¬H) and the initial probability p(H) are uniquely determined from each other in an appropriate way*.

Indeed, the following theorems show that the this possibility does hold.

Theorem 1 For any p(E) and p(E|H) such that p(E) ≠ p(E|H), p(H) depend on p(E|¬H) in the following way:

$$p(H) = \frac{p(E) - p(E|\neg H)}{p(E|H) - p(E|\neg H)}$$

Theorem 1 allows us to study how, for fixed values of p(E) and p(E|H), the initial probability p(H) (= [p(E) – p(E|¬H)]/[p(E|H) – p(E|¬H)]) varies with respect to p(E|¬H).

Theorem 1

For fixed values of p(E) and p(E|H):

(i) if E confirms H, then p(H) decreases as p(E|¬H) increases;

(ii) if E disconfirms H, then p(H) increases as p(E|¬H) increases.

On the basis of Theorem 1, we can prove our main result, i.e., the following equivalence theorem between the L-principles and I-principles:

Theorem 2

(i) N-Ind is equivalent to Imp-Ind.

(ii) N-Dec is equivalent to Imp-Inc.

(iii) N-Inc is equivalent to Imp-Dec.

References

Crupi, Vincenzo. 2015. “Confirmation.” In Stanford Encyclopedia of Philosophy, ed. Edward N. Zalta. Stanford

Festa, Roberto 2012. ‘“For Unto Every One That Hath Shall Be Given”: Matthew Properties for Incremental Confirmation’, *Synthese* **184**, pp. 89–100.

Festa, Roberto & Gustavo Cevolani 2017. ‘Unfolding the Grammar of Bayesian Confirmation: Likelihood and Antilikelihood Principles’, *Philosophy of Science* **84**, no. 1, pp. 56–81.

Kuipers, Theo A. F. 2000. From Instrumentalism to Constructive Realism. Dordrecht: Kluwer.

Roche, William. 2014. "A Note on Confirmation and Matthew Properties." Logic and Philosophy of Science 12:91–101.

ADDENDA FOR DISCUSSION

Addendum 1

It can be proved that the four L-principles above are structural principles (Festa & Cevolani 2017).

Although the above quantitative formulation of some L-principles is appropriate for the exploration of the grammar of confirmation, such principles can be reformulated in comparative terms.

The comparative counterparts of L-Inc, N-Ind, N-Dec, and N-Inc can be stated as follows:

L-Inc* If $p(E|\neg H1) = p(E|\neg H2)$ and $p(E|H1) > p(E|H2)$, then $C(H1,E) > C(H2,E)$.

N-Ind* If $p(E|H1) = p(E|H2)$, then $C(H1,E) = C(H2,E)$.

N-Dec* If $p(E|H1) = p(E|H2)$ and $p(E|\neg H1) > p(E|\neg H2)$, then $C(H1,E) < C(H2,E)$.

N-Inc* If $p(E|H1) = p(E|H2)$ and $p(E|\neg H1) > p(E|\neg H2)$, then $C(H1,E) > C(H2,E)$.

L-Inc*, N-Ind*, N-Dec*, and N-Inc* are identical, or closely related, to the following widely discussed principles:

LL *Law of Likelihood*

(i) If $p(E|H1) = p(E|H2)$, then $C(H1,E) = C(H2,E)$.

(ii) If $p(E|H1) > p(E|H2)$, then $C(H1,E) > C(H2,E)$.

WLL *Weak Law of Likelihood*

If $p(E|H1) > p(E|H2)$ and $p(E|\neg H1) < p(E|\neg H2)$, then $C(H1,E) > C(H2,E)$.

WLL-L *Likelihood ingredient of WLL*

If $p(E|H1) > p(E|H2)$ and $p(E|\neg H1) = p(E|\neg H2)$, then $C(H1,E) > C(H2,E)$.

WLL-N *N-likelihood ingredient of WLL*

If $p(E|H1) = p(E|H2)$ and $p(E|\neg H1) > p(E|\neg H2)$, then $C(H1,E) < C(H2,E)$.

RWLL-N *Reversal of WLL-N*

If $p(E|H1) = p(E|H2)$ and $p(E|\neg H1) > p(E|\neg H2)$, then $C(H1,E) > C(H2,E)$.

Some elucidations on the intuitive content of the principles will be convenient.

LL expresses the intuition that E confirms H1 more than H2 just when H1 is better than H2 in predicting E, i.e., when E is more probable given H1 than given H2.

WLL is one of the several versions of the so-called Weak Law of Likelihood, which owes its name to the fact that such law is a weakening of the Law of Likelihood LL. WLL-L expresses the intuition that if H1 is better than H2 in predicting E and ¬H1 is worse than ¬H2 in predicting E, then E confirms H1 more than H2.

WLL-L and WLL-N are, so to say, the likelihood and N-likelihood ingredients of WLL. Indeed, WLL-L expresses the intuition that if H1 is better than H2 in predicting E and ¬H1 is as good as ¬H2 in predicting E, then E confirms H1 more than H2. Analogously, WLL-N says that if H1 is as good as H2 in predicting E and ¬H1 is worse than ¬H2 in predicting E, then E confirms H1 more than H2.

Finally, RWLL-N is, so to speak, the reversal of principle WLL-N, in the sense that it is obtained from WLL-N by reversing the inequality sign in the second part of the antecedent (hence the label, which stands for “reversed WLL-N”). Accordingly, RWLL-N says that if H1 is as good as H2 in predicting E and ¬H1 is better than ¬H2 in predicting E, then E confirms H1 more than H2.

Note that WLL and WLL-N are widely supported L-principles of Bayesian confirmation while RWLL-N is a very strange and prima facie highly counterintuitive L-principle. For this reason, RWLL-N can be called an *antilikelihood principle*.

However, it has been argued that, quite surprisingly, there are intuitively well-motivated incremental measures that not only violate WLL and WLL-N but indeed satisfy some “antilikelihood principles” that express intuitions diametrically opposed to the ones underlying WLL (Festa & Cevolani 2017).

In particular, it has been argued that some intuitively appealing incremental measures satisfy RWLL-N.

The following logical relations among the above L-principles should be pointed out:

Theorem (i) LL entails WLL.
(ii) LL entails WLL-L.
(iii) WLL-L & WLL-N entails WLL.

One can see that our comparative principles L-IncL*, L-IndN*, L-DecN*, and L-IncN* are closely related to the above L-principles LL, WLL, WLL-L, WLL-N, and RWLL-N. Indeed, L-IndN*, L-DecN*, and L-IncN* are identical to WLL-L, WLL-N, and RWLL-N, respectively.

Addendum 2

Imp-Ind   IS-Ind & IF-Ind

This amounts to saying that the following two clauses hold:

(i) For fixed values of p(E) and p(E|H), with p(E|H) > p(E), when p(H) increases C(H,E) does not change.

(ii) For fixed values of p(E) and p(E|H), with p(E|H) < p(E), when p(H) increases C(H,E) does not change.

The intuitive content of the principles IS-Ind and IF-Ind can be expressed as follows. Recalling that the impact of successes and failures on H changes together with C(H,E), the claim made by IS-Ind and IF-Ind that C(H,E) is not affected by p(H) amounts to the claim that the impact of successes and failures on H is not affected by p(H).

Imp-Inc   IS-Inc & IF-Dec

This amounts to saying that the following two clauses hold:
(i) For fixed values of p(E) and p(E|H), with p(E|H) > p(E), when p(H) increases C(H,E) increases.
(ii) For fixed values of p(E) and p(E|H), with p(E|H) < p(E), when p(H) increases C(H,E) decreases.

The intuitive content of the principles IS-Inc and IS-Dec can be expressed as follows. Recalling that the impact of a success E on H increases when C(H,E) increases, the claim made by IS-Inc that C(H,E) increases when p(H) increases amounts to the claim that the impact of a success on H increases when p(H) increases.

IS-Dec amounts to the claim that the impact of a success on H decreases when p(H) increases.

Imp-Dec   IS- Dec & IF-Inc

This amounts to saying that the following two clauses hold:

(i) For fixed values of p(E) and p(E|H), with p(E|H) > p(E), when p(H) increases C(H,E) decreases.

(ii) For fixed values of p(E) and p(E|H), with p(E|H) < p(E), when p(H) increases C(H,E) increases.

The intuitive content of the principles IF-Inc and IF-Dec can be expressed in terms of impact. Recalling that the impact of a failure E on H decreases when C(H,E) increases, the claim made by IF-Inc that C(H,E) increases when p(H) increases amounts to the claim that the impact of a failure on H decreases when p(H) increases. In a similar way, IF-Dec amounts to the claim that the impact of a failure on H increases when p(H) increases.

Addendum 3

The above I-principles can be reformulated in comparative terms as follows:

I-IncL* If $p(H1) = p(H2)$ and $p(E|H1) > p(E|H2)$, then $C(H1,E) > C(H2,E)$

IS-Ind* If $p(E|H1) = p(E|H2) > p(E)$ and $p(H1) > p(H2)$, then $C(H1,E) = C(H2,E)$

IS-Inc* If $p(E|H1) = p(E|H2) > p(E)$ and $p(H1) > p(H2)$, then $C(H1,E) > C(H2,E)$

IS-Dec* If $p(E|H1) = p(E|H2) > p(E)$ and $p(H1) > p(H2)$, then $C(H1,E) < C(H2,E)$

IF-Ind* If $p(E|H1) = p(E|H2) < p(E)$ and $p(H1) > p(H2)$, then $C(H1,E) = C(H2,E)$

IF-Inc* If $p(E|H1) = p(E|H2) < p(E)$ and $p(H1) > p(H2)$, then $C(H1,E) > C(H2,E)$

IF-Dec* If $p(E|H1) = p(E|H2) < p(E)$ and $p(H1) > p(H2)$, then $C(H1,E) < C(H2,E)$

Imp-Ind* *Impact does not depend on initial probability*
IS-Ind* & IF-Ind*

This amounts to saying that the following two clauses hold:

(i) If $p(E|H1) = p(E|H2) > p(E)$ and $p(H1) > p(H2)$, then $C(H1,E) = C(H2,E)$

(ii) If $p(E|H1) = p(E|H2) < p(E)$ and $p(H1) > p(H2)$, then $C(H1,E) = C(H2,E)$

Imp-Inc* *Impact increases with initial probability*

IS-Inc* & IF-Dec*

This amounts to saying that the following two clauses hold:

(i) If $p(E|H1) = p(E|H2) > p(E)$ and $p(H1) > p(H2)$, then $C(H1,E) > C(H2,E)$

(ii) If $p(E|H1) = p(E|H2) < p(E)$ and $p(H1) > p(H2)$, then $C(H1,E) < C(H2,E)$

Imp-Dec* *Impact decreases with initial probability*

IS- Dec* & IF-Inc*

This amounts to saying that the following two clauses hold:

(i) If $p(E|H1) = p(E|H2) > p(E)$ and $p(H1) > p(H2)$, then $C(H1,E) < C(H2,E)$

(ii) If $p(E|H1) = p(E|H2) < p(E)$ and $p(H1) > p(H2)$, then $C(H1,E) > C(H2,E)$

The above comparative I-principles are identical, or closely related, to some I-principles widely discussed in the literature. More precisely, they are identical, or closely related, to the so-called Matthew principles for Bayesian confirmation studied by Kuipers (2001), Festa (2012), Roche (2014), and Festa and Cevolani (2015).

Let us consider the comparative Imp-principle Imp-Inc*. According to the clause (i) of Imp-Inc* the impact of a success increases when the initial probability of the hypothesis increases. This principle, anticipated by Popper and discussed by Kuipers, was labelled "Matthew effect" by Kuipers, and afterwards "Matthew effect for positive confirmation" by Festa, Festa & Cevolani, and Roche.

According to the clause (ii) of Imp-Inc* the impact of a failure increases when the initial probability of the hypothesis increases. This principle, discussed by Festa, Festa & Cevolani, and Roche was called afterwards "Matthew effect for disconfirmation". The just mentioned authors reserve the label "Matthew effect" to the whole principle Imp-Inc*.

Now let us consider the comparative Imp-principle Imp-Dec*. According to the clause (i) of Imp-Dec* the impact of a success decreases when the initial probability of the hypothesis increases. According to the clause (ii) of Imp-Dec* the impact of a failure decreases when the initial probability of the hypothesis increases. This principle has been discussed by Festa, Festa & Cevolani, and Roche.

The clause Imp-Dec* (i) is called "reverse Matthew effect for positive confirmation" while the clause Imp-Dec* (ii) is called "reverse Matthew effect for disconfirmation." Finally, the whole principle Imp-Dec* is called "reverse Matthew effect".