Luigi Gobbi[1] Last modified 13/May/2009

PDB. Some justifications and some comparisons with Bayesianism

In the 3rd millenium, after the dark Middle Ages (dominated by the *transcendent* God) and after Galileo's heresy, it is very arguable that probability and truth ought to be based on the only evidence. This paper will deal only with a probability of such a conception.

Suppose we live in a world with discrete time and all our long experience consisted in

$$\square, \blacksquare, \square, \blacksquare, \square, \blacksquare, \square, \blacksquare, \square, \blacksquare, \square, \blacksquare, \square, \blacksquare, \ldots, \square, \blacksquare \tag{1}$$

then we all want that our inductive methods will return the next datum of experience to be $\square$ with probability 100%.

Likewise, if we have

$$\square, \square, \square, \square, \square, \square, \square, \square, \ldots, \square, \square \tag{2}$$

then we all want our inductive methods to return the next datum of experience to be $\square$ with *p*robability 100%.

We could indicate this by the writing:

$$\Pi\left(\square, \square, \square, \square, \square, \square, \square, \square, \ldots, \square, \square; \square\right) = 1 \tag{3}$$

Having experienced n data $\square$, then we have got n conclusive verifications of white data. We may have a theory $\mathcal{T}_1$ according to which all data are white; we are commonly led to say that the more verifications of white data we have, the more verified[2] the the theory is. Nonetheless, the data are conclusively verified whereas the theory is non-conclusively verified. However, it would seem intuitive, for this example, to hold that the theory $\mathcal{T}_1$ has been verified n times – *per se*: independently of the other theories we might have conceived for explaining the data of experience and without any reference to prior beliefs. In the case of (2), $\mathcal{T}_1$ would then be verified at every instance, thereby being the most

[1] www.luigigobbi.com

[2] This linguistic phenomenology was in the philosophical use at the time of Ayer, namely with 'weak verification'. After that, what I call 'non-conclusive verification' evolved under the term 'confirmation'. However the one I use is intuitive to the most present-day people, and the notion of confirmation of a theory is quite different from mine as it is based on the probability of the theory given the evidence.

verified theory in comparison with any other theory we conceived. If so, $\mathcal{T}_1$ would become also the most probable (probable 1 in the limit of n infinite) and therefore $\mathcal{T}_1$ would be the only to weigh in evaluating the probability for the next datum (which is the result of (3), as $\Pi^{hy}_{\square\mathcal{T}_i}$ – the *hy*pothetical *p*robability of the future $\square$ according to the theory $\mathcal{T}_1$ – is 1).

More generally, if we face the problem of determining the probability of a next $\square$, after an arbitrary sequence of evidence (and no other hypothesis or information on the evidence available), and having got some consistent theories $\mathcal{T}_i$'s, we could re-formulate the problem of finding an inductive method from how to costrain probabilities to how to costrain the weight to assign to every theory's probability:

$$\Pi_{\square} = \frac{\sum_i \Pi^{hy}_{\square\mathcal{T}_i} \cdot V_{\mathcal{T}_i}}{\sum_j V_{\mathcal{T}_j}} \tag{4}$$

This formulation of the problem is absolutely general so long as there is no constrainig of the weights.

Bayesianism's constraint is:

$$V_{\mathcal{T}_i} = \mathring{\Pi}_{\mathcal{T}_i} \cdot \Pi\left(\mathcal{E}_P|\mathcal{T}_i\right) \tag{5}$$

where $\Pi\left(\mathcal{E}_P|\mathcal{T}_i\right)$ is the likelihood of theory $\mathcal{T}_i$ given the *p*ast *e*vidence $\mathcal{E}_P$ and $\mathring{\Pi}_{\mathcal{T}_i}$ is the prior probability of theory $\mathcal{T}_i$, assigned before the *p*ast *e*vidence $\mathcal{E}_P$.

This Bayesian probability, however, is not solely based on the evidence as there is no unique universal agreement on how to fix the weights.

Let's then look for some desiderata that these weights of the considered probability must fulfil if we want a probability based only on evidence. But can there be a general method for directly quantifying the verifications of every theory $V_{\mathcal{T}_i}$?

Before starting it's useful to introduce some new notation.

After $\Pi^{hy}_{\square\mathcal{T}_i}$, let's introduce $\Pi^{ex}_{\square\mathcal{T}_i}$.

$$\Pi^{ex}_{\square\mathcal{T}_i} = \frac{F^n_{\square\mathcal{T}_i}}{T^n_{\mathcal{T}_i}} \tag{6}$$

is the *ex*perimental probability, or frequency, of the past $\square$ over the instants where $\mathcal{T}_i$ was contemplated. It is equale to the ratio of $F^n_{\square\mathcal{T}_i}$ (*f*avorably and conclusively) verifi*ed* cases of $\square$ over $T^n_{\mathcal{T}_i}$ the *t*otal of (conclusively) verifi*able* cases according to the theory $\mathcal{T}_i$.

This number T will change as the time of the evidence goes by and the superscript t indicates the instant for which T is considered. n is the last instant of the evidence.
The previous likelihood $\Pi\left(\mathcal{E}_P|\mathcal{T}_i\right)$ of (5) is now indicated as $U^n_{\mathcal{T}_i}$. This is done in order to lighten the notation, to make it uniform, and to distinguish this likelihood[3], in the sense that is defined as:[4]

$$U^n_{\mathcal{T}_i} = \eta^n_{\mathcal{T}_i} \cdot \theta^n_{\mathcal{T}_i} = \frac{T^n_{\mathcal{T}_i}!}{\prod\limits_{x\in\sigma}\left(F^n_{x\mathcal{T}_i}!\right)} \cdot \prod_{x\in\sigma}\left(\Pi^{hy}_{x\mathcal{T}_i}\right)^{F^n_{x\mathcal{T}_i}} \tag{7}$$

that is to say, the function η is never omitted – as it is often done in the literature[5].

Then, if we want to base the probability on the sole evidence, it seems definitely reasonable to make the weights of (4) to quantify the "verifications" of each theory by the past evidence, contemplated by each theoretical future probability: in this way we would obtain the intuitive result to have the weights proportional to the verifications of the theory and the probability of a theory proportional to its verifications. This weight verifications V will be function of time and we represent this with $V^t_{\mathcal{T}_i}$ where t indicates the considered instant.

Some desiderata on the $V^t_{\mathcal{T}_i}$ for such a "verificationist" probability are:

1)$V^t_{\mathcal{T}_i}$ should regard the verifications of the only theory $\mathcal{T}_i$ by the evidence $\mathcal{E}_P$ in order to obtain the "right" probability, independently of the other theories that are considered.

2)Obviously we want the attained probability to satisfy the Kolmogorov axioms, when defined. This implies[6] also that $V^t_{\mathcal{T}_i}$ ought to not change for considering different types of evidence when when considering the same theory and future evidence: $V^t_{\mathcal{T}_i}$ ought to be a specific single value on the domain of one theory and one past evidence, independently of the considered future evidence to evaluate.

3)Moreover, as we would like such probability to depend on the evidence solely, the probability must not be defined if no evidence of any kind is available.

4)If $\Pi^{ex}_{\square\mathcal{T}_i} = \Pi^{hy}_{\square\mathcal{T}_i} = 1$ then V is maximum, and in this case $V^t_{\mathcal{T}_i} = T^t_{\mathcal{T}_i} = F^t_{\square\mathcal{T}_i}$: if we have F

[3] Generated by a i.i.d. multinomial distribution, representing the same chance.

[4] x indicates the generic types of piece of evidence. $x = \square$ is one of them. σ is the set of all types of piece of evidence, which may be infinite. U has the mathematical property that $\forall x\ U^n_{\square\mathcal{T}_i} = U^n_{\blacksquare\mathcal{T}_i} = U^n_{x\mathcal{T}_i}$: U is independent of the choice of the particular type of evidence.

[5] As it is said to be a constant that cancel harmlessly. I'm not of the same view; moreover I would replace the factorials $(*)!$ of η by the $\Gamma(1+*)$ function for the case of "fractional" evidence.

[6] As it will be showed later.

favorable and conclusive verifications, as predicted by the theory, then we want the theory to be verified F times.

5)If $0 < \Pi^{ex}_{\Box \mathcal{T}_i} = \Pi^{hy}_{\Box \mathcal{T}_i} < 1$ then $V^t_{\mathcal{T}_i} < T^t_{\mathcal{T}_i}$: the theory is verified (some times), but less than F.

6)If $0 < \Pi^{ex}_{\Box \mathcal{T}_i} < \Pi^{hy}_{\Box \mathcal{T}_i} < 1$ or $0 < \Pi^{ex}_{\Box \mathcal{T}_i} > \Pi^{hy}_{\Box \mathcal{T}_i} < 1$ then $V^t_{\mathcal{T}_i}$ is less than the value of the previous point, for which it is maximum: the theory would be less verified than before.

7)If $\Pi^{ex}_{\Box \mathcal{T}_i} > 0$ and $\Pi^{hy}_{\Box \mathcal{T}_i} = 1$ then $V^t_{\mathcal{T}_i} = 0$: we would have no verifications at all as the theory has been falsified.

8)If $\Pi^{ex}_{\Box \mathcal{T}_i} = \Pi^{hy}_{\Box \mathcal{T}_i}$ then $V^{t_1}_{\mathcal{T}_i} < T^{t_2}_{\mathcal{T}_i}$ when $t_1 < t_2$: the more a theory is verified and the more its V must be.

9)If an essential part of a theory is not verifiable (has not yet received the possibility from evidence of an essential kind to support it), then we couldn't consider the whole theory as verifiable.

10)The attained probability ought not to be affected by "skeptical paradoxes" such as the Goodman's new riddle of induction and by other types of paradoxes, thereby solving the problem of induction.

11)And, on this line, why not, also to give a general account of vagueness and causality and more.

By endorsing some "empiricist intuitions", I propose the following prescriptive principles for a "truly verificationist" probability:

1] $V^t_{\mathcal{T}_i}$ ought to be proportional to the likelihood of its theory: $V^t_{\mathcal{T}_i} \propto U^t_{\mathcal{T}_i}$ (the more likely, the more probable)

2] $V^t_{\mathcal{T}_i}$ ought to be proportional to the absolute frequency of data on which the likelihood of the theory $\mathcal{T}_i$ has been computed: $V^t_{\mathcal{T}_i} \propto T^t_{\mathcal{T}_i}$ (the more verifiable, the more probable)

By setting

$$V^t_{\mathcal{T}_i} = U^t_{\mathcal{T}_i} \cdot T^t_{\mathcal{T}_i} \tag{8}$$

the *prescriptive* principles are fulfilled. I call this obtained quantity as '*p*rescriptive *v*erifications' of a theory (or PV for short).

The more we test a theory by further evidence, the more the verifications keep on varying[7] by time, therefore these verifications are clearly *non conclusive*.

The first prescriptive principle is, more or less explicitly, held by Bayesianism and Likelihoodism; however this principle alone doesn't even suggest that the probability gets "better" by having more evidence.

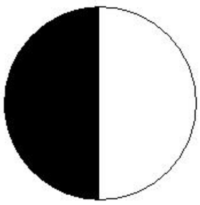
The second prescriptive principle is intuitive too if we consider that we want to take a larger piece of evidence as more "*t*ransparent" and more representative of the chance that produced it. The more evidence of the theory we have, the more the theory ought to be weighted.

My claim is that Bayesianism (with its choice of the weight) fulful only a few of the listed desiderata, but my choice fulfils every of them.

In this summary I show some practical numerical examples for the first 10 desiderata.

Let's start by computing the verifications V of some numerical example to see if our desiderata are fulfilled and we start with one of the most elementary example:

let's consider the theory according to which there is chance $\frac{1}{2}$ to produce the data. Let's call this theory $\mathcal{T}_C$[8]. Using a visual example, the data of the evidence could be the outcome of the spinning of the following "wheel of fortune" χ_7^{hy} of `Figure 1`



`Figure 1`

Moreover suppose all our past evidence consists in:

$$I_{20} = \square, \blacksquare, \square, \blacksquare, \square, \blacksquare, \square, \blacksquare, \square, \blacksquare, \square, \blacksquare, \square, \blacksquare, \square, \blacksquare, \square, \blacksquare, \square, \blacksquare,$$
$$\square, \blacksquare, \square, \blacksquare, \square, \blacksquare, \square, \blacksquare, \square, \blacksquare, \square, \blacksquare, \square, \blacksquare, \square, \blacksquare, \square, \blacksquare, \square, \blacksquare \quad (9)$$

Then, if we are to determine the probability of a next $\square$ given I_{20}, we can now compute

[7] Unless the theory is falsified.

[8] Such as 'theory of the (unbiased) *c*oin.'

the values of `Table 1`.[9]

F	$T-F$	U	V
0	0	1	0
1	1	0.5	1
2	2	0.375	1.5
3	3	0.312	1.875
4	4	0.273	2.187
5	5	0.246	2.461
6	6	0.225	2.707
7	7	0.209	2.932
8	8	0.196	3.142
9	9	0.185	3.338
10	10	0.176	3.524
11	11	0.168	3.700
12	12	0.161	3.868
13	13	0.155	4.030
14	14	0.149	4.184
15	15	0.144	4.334
16	16	0.140	4.478
17	17	0.136	6.618
18	18	0.132	4.754
19	19	0.129	4.886
20	20	0.125	5.014

`Table 1`

So, it turns out that the weight of $\Pi_{\Box\mathcal{T}_C}$ is 5.014. The desideratum 5 is satisfied as $0 < 5.014 < T^{40}_{\mathcal{T}_C} = 40$.

The desideratum 8 is satisfied too, in fact, the verifications of the theory $\mathcal{T}_C$ by $\mathcal{E}_P$ are strictly increasing over time[10]. It is curious to report that I asked some university teachers

[9] In the following tables the numbers are rounded to the third decimal place.

[10] I can show this by recalling that

$$V = T \cdot U = T \cdot \eta \cdot \theta \tag{10}$$

and a formula proved by de Moivre:

$$\eta = \left(\begin{array}{c} n \\ F^n_\Box \end{array} \right) \sim \frac{2^{n+1}}{\sqrt{2\pi n}} e^{-\frac{(n-2F^n_\Box)^2}{2n}} \tag{11}$$

Then, for the evidence I_n where n is the even length of the string (so that $\Pi^{ex}_{\Box\mathcal{T}_C} = \Pi^{hy}_{\Box\mathcal{T}_C}$):

$$\eta^n_{\mathcal{T}_C} = \left(\begin{array}{c} n \\ n/2 \end{array} \right) \sim \frac{2^{n+1}}{\sqrt{2\pi n}} \tag{12}$$

and obviously

$$\theta^n_{\mathcal{T}_C} = \frac{1}{2^n} \tag{13}$$

So we can write

$$U^n_{\mathcal{T}_C} \sim \frac{2^{n+1}}{\sqrt{2\pi n}} \frac{1}{2^n} \tag{14}$$

and finally

$$V^n_{\mathcal{T}_C} \sim \frac{2^{n+1}}{\sqrt{2\pi n}} \frac{1}{2^n} n = \sqrt{\frac{2n}{\pi}} \tag{15}$$

shows that $V^n_{\mathcal{T}_C}$ is strictly increasing over *all* time.

of mathematical subjects whether a likelihood multiplied by a first-power term could have approached infinity, and they all told me no: as usual, the likelihood is confused with the only θ function which is a decreasing exponential, therefore a power "can't do much" in comparison; however my philosophical intuitions proved to be stronger than the mathematical ones.

Here I show also some cases similar to the previous one, where there is an evidence and a theory such that $\Pi^{ex}_{\square\mathcal{T}_i} = \Pi^{hy}_{\square\mathcal{T}_i} = \frac{3}{4}$:

F	$T-F$	U	V
0	0	1	0
3	1	0.422	1.687
6	2	0.311	2.492
9	3	0.258	3.097
12	4	0.225	3.603
15	5	0.202	4.047
18	6	0.185	4.447
21	7	0.172	4.812
24	8	0.161	5.153
27	9	0.152	5.472
30	10	0.144	5.774

Table 2

and $\Pi^{ex}_{\square\mathcal{T}_i} = \Pi^{hy}_{\square\mathcal{T}_i} = \frac{9}{10}$:

F	$T-F$	U	V
0	0	1	0
9	1	0.387	3.874
18	2	0.285	5.073
27	3	0.236	7.082
36	4	0.206	8.236

Table 3

This last table could be used also to show the complete verificationist inappropriateness of the Bayesian weight: evaluating a theory of a theory with chance $\frac{9}{10}$ by the evidence

$$L_{36} = \square, \blacksquare, \square, \square, \square, \square, \square, \square, \square, \square, \square, \blacksquare, \square, \square, \square, \square, \square, \square, \square, \square,$$
$$\square, \blacksquare, \square, \square, \square, \square, \square, \square, \square, \square, \square, \blacksquare, \square, \square, \square, \square, \square, \square, \square, \square \quad (16)$$

the Bayesian weight at $T = 40$ is practically the same[11] as the Bayesian weight for the same problem if there were just the first two pieces of evidence: $I_1 = \square, \blacksquare$ – at $T = 2$. In

[11]The Bayesian weight is equal to $\mathring{\Pi}_{\mathcal{T}_i} \cdot U^n_{\mathcal{T}_i}$. The prior is a constant, therefore if U is the same then the Bayesian weight is the same.

fact U would here be $0.18 = 2 \cdot 0.9 \cdot 0.1$. This goes too much against the intuition to be acceptable!! By contrast, for I_1 my weight $V = UT$ would intuitively be only $0.18 \cdot 2 = 0.36$ – which is about 23 times smaller than the weight for $T = 40$!

Another example can be done by the `Table 1`: the Bayesian weight of a theory with chance $\frac{9}{10}$ by the evidence I_{20} would be at $T = 40$ the same as the one of the same theory by the evidence $B_3 = \blacksquare, \blacksquare, \blacksquare$ ($F = 0$, $T - F = 3$, $U = 0.125$). By constrast, my weight PV would be $0.125 \cdot 3 = 0.375$ which is about 13 times smaller than the PV by I_{20} at $T = 40$; $\mathcal{T}_C$ has very few "absolute" verifications in the case of B_3 and it sensibly is 13 times more verified by I_{20}.

The Bayesian weight cannot make any difference on whether it is low (or it has a value in general) because of a large evidence or because of a "bad fitting" between evidence and theory. In fact, many world expert of applied probability theory complain about the lack of a tool to divide cases of intrinsic randomness of the data from an insufficient data information. In my account the ignorance on data is removed when we have a lot of verifications, literally; what possibly remains is just intrinsic randomness of the data.

Once again, the tables showed the only instants where the desiderata 5 and 8 are concerned ($\Pi^{ex}_{\square \mathcal{T}_i} = \Pi^{hy}_{\square \mathcal{T}_i}$), but, in general and obviously, the verifications are don't always increase by any further evidence and can decrease:

F	$T - F$	U	V
9	1	0.387	3.874
9	2	0.213	2.344
10	1	0.384	4.219
18	2	0.285	5.073

`Table 4`

Some more cases where the evidence and the theory are such that $\Pi^{ex}_{\square \mathcal{T}_i} = \Pi^{hy}_{\square \mathcal{T}_i} = \frac{39}{40}$:

F	$T - F$	U	V
0	0	1	0
39	1	0.373	14.901

`Table 5`

And $\Pi^{ex}_{\square \mathcal{T}_i} = \Pi^{hy}_{\square \mathcal{T}_i} = \frac{40}{40}$:

F	$T - F$	U	V
0	0	1	0
40	0	1	40

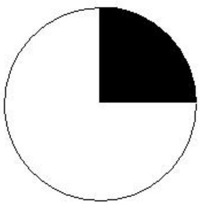
`Table 6`

This last table shows that the desideratum 4 is satisfied as well.

So far we have showed the values of $\Pi^{ex}_{\square\mathcal{T}_i} = \Pi^{hy}_{\square\mathcal{T}_i}$ greater than $\frac{1}{2}$, but if we want to determine the U's and V's of the cases smaller than $\frac{1}{2}$, then they comply with the property of being specular with respect to $\frac{1}{2}$. In fact I can make this clear by a visual example referred to `Table 2`: an evidence compatible with it is:

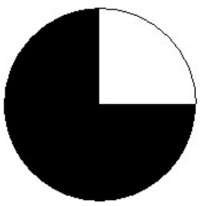
$$G_{30} = \square, \blacksquare, \square, \square, \square, \blacksquare, \square, \square, \square, \blacksquare, \square, \square, \square, \blacksquare, \square, \square, \square, \blacksquare, \square, \square,$$
$$\square, \blacksquare, \square, \square, \square, \blacksquare, \square, \square, \square, \blacksquare, \square, \square, \square, \blacksquare, \square, \square, \square, \blacksquare, \square, \square \qquad (17)$$

The theory of chance $\frac{3}{4}$ for $\square$ that generated (17) could simply be the wheel χ^{hy}_{10} of `Figure 2`



`Figure 2`

If we now want to consider a theory of chance $\frac{1}{4}$ for $\square$ such as χ^{hy}_{9} of `Figure 3`



`Figure 3`

has the following evidence compatible with having $\Pi^{ex}_{\square\mathcal{T}_i} = \Pi^{hy}_{\square\mathcal{T}_i}$:

$$H_{10} = \blacksquare, \square, \blacksquare, \blacksquare, \blacksquare, \square, \blacksquare, \blacksquare, \blacksquare, \square, \blacksquare, \blacksquare, \blacksquare, \square, \blacksquare, \blacksquare, \blacksquare, \square, \blacksquare, \blacksquare,$$
$$\blacksquare, \square, \blacksquare, \blacksquare, \blacksquare, \square, \blacksquare, \blacksquare, \blacksquare, \square, \blacksquare, \blacksquare, \blacksquare, \square, \blacksquare, \blacksquare, \blacksquare, \square, \blacksquare, \blacksquare \qquad (18)$$

Therefore it is evident that χ^{hy}_{10} is verified by G_{30} exactly as much as χ^{hy}_{9} is verified by H_{10} for the only color has been exchange between the two cases.

Anyway here are two numerical examples too:

F	$T-F$	U	V
0	0	1	0
1	39	0.373	14.901

Table 7

F	$T-F$	U	V
0	0	1	0
0	40	1	40

Table 8

At this point, after noting that $0 = \frac{0}{40}$, $\frac{1}{40} = \frac{1}{40}$, $\frac{1}{10} = \frac{4}{40}$, $\frac{1}{4} = \frac{10}{40}$, $\frac{1}{2} = \frac{20}{40}$, $\frac{3}{4} = \frac{30}{40}$, $\frac{9}{10} = \frac{36}{40}$, $\frac{39}{40} = \frac{39}{40}$ and $1 = \frac{40}{40}$ we can write[12]

$\Pi_{\square}^{ex} = \Pi_{\square}^{hy} = \frac{0}{40}$	$\Pi_{\square}^{ex} = \Pi_{\square}^{hy} = \frac{1}{40}$	$\Pi_{\square}^{ex} = \Pi_{\square}^{hy} = \frac{4}{40}$	$\Pi_{\square}^{ex} = \Pi_{\square}^{hy} = \frac{10}{40}$
$V = 40$	$V = 14.90$	$V = 8.23$	$V = 5.77$

$\Pi_{\square}^{ex} = \Pi_{\square}^{hy} = \frac{20}{40}$	$\Pi_{\square}^{ex} = \Pi_{\square}^{hy} = \frac{30}{40}$	$\Pi_{\square}^{ex} = \Pi_{\square}^{hy} = \frac{36}{40}$	$\Pi_{\square}^{ex} = \Pi_{\square}^{hy} = \frac{39}{40}$	$\Pi_{\square}^{ex} = \Pi_{\square}^{hy} = \frac{40}{40}$
$V = 5.01$	$V = 5.77$	$V = 8.23$	$V = 14.90$	$V = 40$

Table 9

which shows that V is minimum for $\Pi_{\square}^{ex} = \Pi_{\square}^{hy} = \frac{1}{2}$. This too is expected as the proportion between the two colors becomes just one in the limit, where we have the maximum of verifications – equal to the number of verifiable cases; the minimum therefore is when the proportion is "opposite" to the limit.

I report now the result of the computation of the verifications of an evidence of 40 verifiable cases ($T = 40$) where $\Pi_{\square}^{ex} = \frac{10}{40}$ (such as H_{10}), but $\Pi_{\square}^{hy}$ varies:

$\Pi_{\square}^{hy} = \frac{0}{40}$	$\Pi_{\square}^{hy} = \frac{1}{40}$	$\Pi_{\square}^{hy} = \frac{4}{40}$	$\Pi_{\square}^{hy} = \frac{10}{40}$
$V = 0$	$V = 1.51 \cdot 10^{-6}$	$V = 1.44 \cdot 10^{-1}$	$V = 5.77$

$\Pi_{\square}^{hy} = \frac{20}{40}$	$\Pi_{\square}^{hy} = \frac{30}{40}$	$\Pi_{\square}^{hy} = \frac{36}{40}$	$\Pi_{\square}^{hy} = \frac{39}{40}$	$\Pi_{\square}^{hy} = \frac{40}{40}$
$V = 3.08 \cdot 10^{-2}$	$V = 1.65 \cdot 10^{-9}$	$V = 1.18 \cdot 10^{-20}$	$V = 2.28 \cdot 10^{-38}$	$V = 0$

Table 10

I display also the verifications at $T = 40$ of a theory that has chance $\Pi_{\square}^{hy} = \frac{1}{4} = \frac{10}{40}$ (such as the wheel of Figure 3), but the $F_{\square}^{40}$'s varies, so that $\Pi_{\square}^{ex}$ varies[13]:

$\Pi_{\square}^{ex} = \frac{0}{40}$	$\Pi_{\square}^{ex} = \frac{1}{40}$	$\Pi_{\square}^{ex} = \frac{4}{40}$	$\Pi_{\square}^{ex} = \frac{10}{40}$
$V = 4.02 \cdot 10^{-4}$	$V = 4.00 \cdot 10^{-3}$	$V = 4.53 \cdot 10^{-1}$	$V = 5.77$

$\Pi_{\square}^{ex} = \frac{20}{40}$	$\Pi_{\square}^{ex} = \frac{30}{40}$	$\Pi_{\square}^{ex} = \frac{36}{40}$	$\Pi_{\square}^{ex} = \frac{39}{40}$	$\Pi_{\square}^{ex} = \frac{40}{40}$
$V = 1.59 \cdot 10^{-2}$	$V = 1.65 \cdot 10^{-9}$	$V = 2.45 \cdot 10^{-16}$	$V = 3.97 \cdot 10^{-21}$	$V = 3.30 \cdot 10^{-23}$

[12] In the following tables the numbers are rounded to the second decimal place.
[13] Here the numerator of the fraction of $\Pi_{\square}^{ex}$ stands for the number F.

Table 11

These last two tables show that desiderata 6 and 7 are fulfilled: V is maximum just for $\Pi_{\square}^{ex} = \Pi_{\square}^{hy} = \frac{1}{4}$ and the theory is falsified by an evidence such as H_{10} when the theoretical chance is 0 or 1.[14]

The Bayesian weights don't fulfil desideratum 4 as they can't exceed 1; they fulfil desiderata 5 and 6 – tough improperly for the same reason; they fulfil[15] desideratum 7; they don't fulfil desideratum 8 – as the tables have extensively showed.

Another remarkable difference is that the Bayesian weight is in general larger than zero before the evidence, but my verificationist weight is 0 if no evidence has come yet.

Let's now make a numerical example to calculate the probability from the verifications of the theories by evidence.

As an evidence[16] we take the (discrete) temporal sequence $\mathcal{E}$ of $n = 4$ data

$\mathcal{E} = \square, \blacksquare, \square, \blacksquare$

We happen –because of any reason– also to have four candidate theories to explain this evidence and predict new one (a $\square$ in this example):

$\mathcal{T}_1$ according to which all the colors are white only

$\mathcal{T}_2$ according to which there is always white and black alternately (with white first)

$\mathcal{T}_3$ according to which the colors are given by the tossing of a fair coin: whites for heads and blacks for tails

$\mathcal{T}_4$ according to which all the colors are black only.

---

[14]Before I proved the fulfilment of desideratum 8 for the only $\mathcal{T}_C$. I showed numerical examples that other theories for which $\Pi_{\square}^{ex} = \Pi_{\square}^{hy} \neq \frac{1}{2}$ (I call the generic theory related to that chance as $\mathcal{T}_{\not{C}}$: *biased coin*) are strictly increasing too, but I haven't proved this in general.
By using again the de Moivre's approximation, we have proved the generality if the disequality

$$V_{\mathcal{T}_{\not{C}}}^{n} \sim n\sqrt{\frac{2}{\pi n}}2^{n} e^{-\frac{(n-2F_{\square}^{n})^2}{2n}} \left(\frac{F_{\square}^{n}}{n}\right)^{F_{\square}^{n}} \left(\frac{n-F_{\square}^{n}}{n}\right)^{n-F_{\square}^{n}} > \sqrt{\frac{2n}{\pi}} \sim V_{\mathcal{T}_C}^{n} \qquad (19)$$

holds true for $0 \leq F_{\square}^{n} \leq n$ and $F_{\square}^{n} \neq \frac{n}{2}$.
But proving the previous disequality true is equivalent to proving the following disequality true:

$$2^{n}\ (\Pi_{\square}^{ex})^{n\Pi_{\square}^{ex}}\ (1-\Pi_{\square}^{ex})^{n(1-\Pi_{\square}^{ex})} - e^{\frac{n}{2}(1-2\Pi_{\square}^{ex})^2} > 0 \qquad (20)$$

which has minimum for $\Pi_{\square}^{ex} = \frac{1}{2}$.

[15]Hoewever I've never seen the general number of falsifications of a theory from Bayesianism, while I give it as a natural mathematical quantification from the likelihood (considering also the case of fractional evidence).

[16]'Evidence' derives from Latin 'evidens', which means 'visible, evident', as it is for this example with colors.

Every theory can be easily seen as an allocation of wheels of fortune and the temporal positions: for example

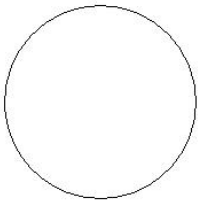


Figure 4

the wheel χ_4^{hy} of Figure 4 can be assigned to the temporal positions $\lambda_1^{hy} = 1, 2, 3, \dots$ and spinned at every instant for generating the datum of experience corresponding to the color pointed by the indicator under which the wheel has stopped. In this way the theory $\mathcal{T}_1$ can be seen as the ordered couple $\left\{ \left(\lambda_1^{hy}, \chi_4^{hy} \right) \right\}$.

Similarly, $\mathcal{T}_4$ can be seen as $\left\{ \left(\lambda_1^{hy}, \chi_8^{hy} \right) \right\}$ with χ_8^{hy} of Figure 5

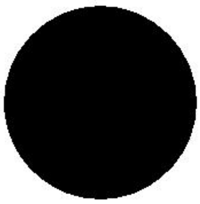


Figure 5

$\mathcal{T}_3$ as $\left\{ \left(\lambda_1^{hy}, \chi_7^{hy} \right) \right\}$ with χ_7^{hy} of Figure 1.

And $\mathcal{T}_2$ as $\left\{ \left(\lambda_2^{hy}, \chi_4^{hy} \right), \left(\lambda_3^{hy}, \chi_8^{hy} \right) \right\}$ with $\lambda_2^{hy} = 1, 3, 5, 7, \dots$ and $\lambda_3^{hy} = 2, 4, 6, 8, \dots$. The wheels can uniquely represent the *chance* specified at every instant according to each theory.

In this case we obtain $V_{\mathcal{T}_1}^4 = 0 \cdot 4 = 0$, $V_{\mathcal{T}_2}^4 = 1 \cdot 4 = 4$, $V_{\mathcal{T}_3}^4 = \frac{(2+2)!}{2!2!} \cdot \left(\left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^2 \right) \cdot 4 = \frac{3}{8} \cdot 4 = 1.5$ and $V_{\mathcal{T}_4}^4 = 0 \cdot 4 = 0$, satisfying our constraints.

Therefore we have all what is needed to compute the (4): $\Pi_{\square}^4 = 0.7\overline{72}$.

As the project seems to have succeeded, it also absolutely natural to define the probability of every theory as the its relative PV's:[17]

$$\Pi_{\mathcal{T}_i} = \frac{V_{\mathcal{T}_i}}{\sum_j V_{\mathcal{T}_j}} \tag{22}$$

Here we would get $\Pi^4_{\mathcal{T}_1} = 0/5.5 = 0$, $\Pi^4_{\mathcal{T}_2} = 4/5.5 = 0.\overline{72}$, $\Pi^4_{\mathcal{T}_3} = 1.5/5.5 = 0.\overline{27}$, $\Pi^4_{\mathcal{T}_4} = 0/5.5 = 0$.

I call the probability based on my prescriptive principles as *PDC*: *p*rescriptive *d*egree of *c*ertainty.

There is however a fundamental important remark to be done: the two prescriptive principles can be well justified only if the data are produced by the "same wheel": i.i.d. hypothesis of multinomial distribution[18]. Therefore all that has been said so far holds only for these "elementary cases of theories", which we'll call *CMS*-rules ρ (elementary with respect to a *conclusive verification.* The (8) must therefore be revised as

$$V^n_\rho = U^n_\rho \cdot T^n_\rho \tag{23}$$

Nevertheless, a rule is nothing but the ordered couple that has just been mentioned.
So, knowing how to derive the various V_ρ's,. the problem is how to derive $V_{\mathcal{T}_i}$ from the V_ρ's in the general case.
In the problem above, theory $\mathcal{T}_2$ was made of two rules V_ρ's, both of value 2: $V_\rho = \frac{2!}{2!0!} \cdot 1^2 0^0 \cdot 2 = 2$.
This example suggests, also for desideratum 4, that it seems right to hold as a general $V_{\mathcal{T}_i}$ the sum of the verification of its rules:

$$V^n_{\mathcal{T}_i} = \bigoplus_{\rho \in \mathcal{T}_i} V^n_\rho = \begin{cases} \sum\limits_{\rho \in \mathcal{T}_i} V^n_\rho & \text{if} \quad \prod\limits_{\rho \in \mathcal{T}_i} V^n_\rho > 0 \\ 0 \quad \text{if} \quad \prod\limits_{\rho \in \mathcal{T}_i} V^n_\rho = 0 \end{cases} \tag{24}$$

The previous formula expresses the third prescriptive principle of PV and PDB, and it contains also an important particularity:

[17] It might be relevant to note that Bayesianism does the very same move. Moreover it follows also the remarkable equation

$$\Pi_{\mathcal{E}_F} = \sum_i \Pi_{\mathcal{E}_F \mathcal{T}_i} \cdot \Pi_{\mathcal{T}_i} \tag{21}$$

[18] For example: if the likelihood of the theory changes through the evidence, which likelihood to consider? Or suppose we consider a value $\Pi^{hy} = \Pi^{ex}$, then the more T we have, the more it is sensible to hold the previous equation; but not if the wheel that "produces" the data is exchanged with another one.

3]The verifications of a theory amount to the sum of the verifications of its rules, unless a theory contains a rule that has no verifications by evidence at all, in which the whole theory must be taken with 0 verifications[19].

In this sense the rules are essential parts of a theory – with respect to its verification: if a “wheel” has received no evidence in its support, then it must be taken as not verified and so must any theory containing it.

The remarkable case of 0 verifications of a rule occurs either when a rule of the theory is falsified ($U = 0$) or when a rule of the theory is not yet verifiable, therefore it is sensible to take the whole theory as falsified or not verifiable (desideratum 9).

Another and last generalization of the prescriptive principles occurs considering that the same wheel could be associated to different causes, therefore a rule should be defined as the ordered triple $\left(\mathcal{C}_\rho^{hy}, \lambda_\rho^{hy}, \chi_\rho^{hy}\right)$, $\mathcal{C}_\rho^{hy}$ being its cause: the whole theory cannot be taken for verified if an essential part of it isn't.

All 3 prescriptive principles are strongly empiricist: the probability of a theory depends solely on the verifications of the evidence.

As a consequence the probability is not defined if no evidence is available (desideratum 3):

$$\Pi_{\mathcal{T}_i} = \frac{V_{\mathcal{T}_i}}{\sum_j V_{\mathcal{T}_j}} = \frac{0}{0} \tag{25}$$

However, when the probability is defined (i.e., when some theory has some verifications, by evidence) then:

$$\sum_{\mathcal{E}_F \in \sigma} \Pi_{\mathcal{E}_F} = \sum_{\mathcal{E}_F \in \sigma} \frac{\sum\limits_{\mathcal{T}_i \in \mathbf{T}} \Pi^{hy}_{\mathcal{E}_F \mathcal{T}_i} \cdot V^n_{\mathcal{T}_i}}{\sum\limits_{\mathcal{T}_i \in \mathbf{T}} V^n_{\mathcal{T}_i}} = \frac{\sum\limits_{\mathcal{T}_i \in \mathbf{T}} \sum\limits_{\mathcal{E}_F \in \sigma} \Pi^{hy}_{\mathcal{E}_F \mathcal{T}_i} \cdot V^n_{\mathcal{T}_i}}{\sum\limits_{\mathcal{T}_i \in \mathbf{T}} V^n_{\mathcal{T}_i}} \tag{26}$$

$$= \frac{\sum\limits_{\mathcal{T}_i \in \mathbf{T}} \left(V^n_{\mathcal{T}_i} \sum\limits_{\mathcal{E}_F \in \sigma} \Pi^{hy}_{\mathcal{E}_F \mathcal{T}_i} \right)}{\sum\limits_{\mathcal{T}_i \in \mathbf{T}} V^n_{\mathcal{T}_i}} = \frac{\sum\limits_{\mathcal{T}_i \in \mathbf{T}} V^n_{\mathcal{T}_i}}{\sum\limits_{\mathcal{T}_i \in \mathbf{T}} V^n_{\mathcal{T}_i}} = 1 \tag{27}$$

it fulfils the Kolmogorov axioms (desideratum 2).

Bayesianism doesn't fulfil desiderata 4 and 9 I showed a problematic case for 2 as well: the step between (26) and (27) was allowed just because $V^n_{\mathcal{T}_i}$ was taken to be independent

[19]But it would be possible to consider a “sub-theory” as verified (more than 0).

of $\mathcal{E}_F$; my $V^n_{\mathcal{T}_i}$ is mathematically proven to be so, but Bayesianism requests just coherence on the priors and this couldn't amount to independency over the type of evidence[20].

I now continue the previous example to elucidate two important observations: in my account new theories can be added without having to modify the weights of the other theories (such as in Bayesianism) – desideratum 1.

So let's suppose we have a new theory $\mathcal{T}_5$, according to which there is always white at the odd instants, and we want to establish $\Pi^4_{\square}$. $\mathcal{T}_5 = \left\{\left(\mathcal{C}^{hy}_1, \lambda^{hy}_2, \chi^{hy}_4\right)\right\}$ where we take $\mathcal{C}^{hy}_1$ to be any cause associated to only $\mathcal{T}_5$. Then we would have a new $\Pi^4_{\square}$ by considering also $\mathcal{T}_5$. However, without considering $\mathcal{T}_5$ the Bayesian probability and *PDB* would be numerically equivalent: in many cases treated by textbooks, predictions using Bayesianism and *PDB* coincide. This seems to happen whenever the Bayesian priors of the theory are proportional to the T's of the theory[21], but this is not the case when $\mathcal{T}_5$ is considered as $T^4_{\mathcal{T}_5} = 2 \neq 4$. To be more precise, the two formulation of the weights would be equivalent if the priors are the same for every theory (hypothesis of uniform distribution on the priors). But then the priors could go out of the sum and cancel with one another: it would amount to considering only the likelihoods and I can't hold the equivalent, i.e., independently of the T's on which the likelihood have been computed.[22]

Another example: let's introduce the theory $\mathcal{T}_6 = \left\{\rho = 61 = \left(\lambda^{hy}_2, \chi^{hy}_7\right), \rho = 62 = \left(\lambda^{hy}_3, \chi^{hy}_{10}\right)\right\}$: fair coin for the odd instants and the tetrahedral dice in the even instants. The determination of $V^4_{\mathcal{T}_6}$ would require the third prescriptive principles in my account: we need to "o-sum" the verifications of the two "elementary subtheories":

$$V^4_{\mathcal{T}_6} = V^4_{61} \oplus V^4_{62} = 2 \cdot \frac{2!}{2!0!}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^0 \oplus 2 \cdot \frac{2!}{0!2!}\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)^0 = \frac{1}{2} \oplus \frac{9}{8} = \frac{13}{8} \quad (28)$$

I'm not sure about the exact way that Bayesians would determine the Bayesian weight of for $\Pi^{hy}_{\square \mathcal{T}_7}$, but I've never found anything analogous to the third prescriprive principle of *PDB* in Bayesianism[23]. Therefore Bayesian and *PDB* predictions wouldn't coincide in

[20]See my fourth argument in my *Sixteen arguments against Bayesiansim.*

[21]And whenever the Bayesian likelihoods are proportional to the U's.

[22]Another difference is that my account must consider all the experienced evidence, while Bayesianism can only consider a "last" part of received evidence by means of a "lastly-obtained" prior. I have a quite critical argument against the Bayesian updating: see my sixth argument against Bayesianism.

[23]Actually the third principle would be forbidden in Bayesianism as it could return that the peobability

general even on cases of the same T's.

Now a note about the desideratum 10:
unlike Bayesianism, the theory of PDB uniquely determines the probability of a theory given some evidence. This is not true however: it is true for the verifications of the theory, but not for its probability which will depend on the set $\mathbf{T}$ of all the considered theories. This feature is common with Bayesianism, however Bayesianism doesn't solve the problem of induction for its incapability at universally fixing the priors[24]. Moreover Bayesianism can't really solve paradoxes such as the 'grue-emeralds' paradox.[25]
The 'grue-emeralds' hypothesis is such that all emeralds are green until a future instant, but blue after that. The usual conclusion is that evidence of green emeralds supports the conclusion that the emeralds will be blue in future.
However this sort of "skeptical" hypothesis receives 0 verifications the theory of PDB as it is made of two rules, the second of which has never been verified![26]

Another huge philosophical difference between Bayesianism and the theory of PDB is that the verifications of a theory can sensibly constitute the fundamental bricket for an monist methodology which encompasses truth too, which would then be regarded as conclusive verification. Methodology which would comfortably apt to facing the most variegated types of methodological problems.
Moreover the PV is not really one of the many confirmation measure as in that case those are founded on the probability, while PV founds the probability.

I conclude by giving a little taste for the desideratum 11.

of some evidence is higher than 1.

[24]The washing out of the priors is not important: the right way for making a prediction should be available even on the basis of little evidence.

[25]Another usual paradox of induction is the Hempel's paradox of black $r$avens that could be tackled by

$$\mathbf{R} = \overset{\bullet}{N_1}, \quad \overset{\circ}{N_2}, \quad \overset{\bullet}{R_3}, \quad \ldots, \quad \overset{\circ}{N_{n-3}}, \quad \overset{\bullet}{R_{n-2}}, \quad \overset{\circ}{N_{n-1}}, \quad \overset{\circ}{N_n}; \quad \overset{\bullet}{R_{n+1}}, \quad \ldots \tag{29}$$

[26]The lack of PV's for a rule of a hypothesis, despite their *possibility*, could actually serve as definition of skeptical hypotheses.

Post scriptum: all cases for the probability and truth of a theory $\mathcal{T}_i$ in *CMS*:

$V^t_{\mathcal{T}_i}$: verifications V of a theory $\mathcal{T}_i$ by evidence $\mathcal{E}^t$ (at a certain time and for the theoretical context)

$F^t_{\mathcal{T}_i}$: falsifications F of a theory $\mathcal{T}_i$ by evidence $\mathcal{E}^t$ (at a certain time and for the theoretical context)

$T^t_{\mathcal{T}_i}$: number of pieces of evidence, parts of $\mathcal{E}^t$, on which the likelihood U of a theory $\mathcal{T}_i$ was contemplated (at a certain time and for the theoretical context)

1] $V^t_{\mathcal{T}_i} > 0$ due to "verifiability" $T^t_{\mathcal{T}_i} > 0$ and "likelihood" $U^t_{\mathcal{T}_i} > 0$:

a) $V^t_{\mathcal{T}_i} > 0$ and every other theory has $V^t_{\mathcal{T}_{i \neq j}} << V^t_{\mathcal{T}_i} \Longrightarrow \Pi^t_{\mathcal{T}_i} = 1$. If $T^t_{\mathcal{T}_i} >> 0$ (else $\mathcal{T}_i$ is not a truth-bearer) $\Longrightarrow \mathcal{T}_i$ is true ($\underline{V}^t_{\mathcal{T}_i} = 1$) ...and if $T^t_{\mathcal{T}_i} >>> 0 \Longrightarrow \mathcal{T}_i$ is very true (and truer)

b) $V^t_{\mathcal{T}_i} > 0$ and another theory has $V^t_{\mathcal{T}_{i \neq j}}$ comparable with $V^t_{\mathcal{T}_i} \Longrightarrow 0 < \Pi^t_{\mathcal{T}_i} < 1$. If $T^t_{\mathcal{T}_i} >> 0$ (else $\mathcal{T}_i$ is not a truth-bearer) $\Longrightarrow \mathcal{T}_i$ has intermediate degree of truth ($0 < \underline{V}^t_{\mathcal{T}_i} = \Pi^t_{\mathcal{T}_i} < 1$)

c) $V^t_{\mathcal{T}_i} > 0$ and no other theory has $V^t_{\mathcal{T}_{i \neq j}} > 0 \Longrightarrow \Pi^t_{\mathcal{T}_i} = 1$. If $T^t_{\mathcal{T}_i} >> 0$ (else $\mathcal{T}_i$ is not a truth-bearer) $\Longrightarrow \mathcal{T}_i$ is true ($\underline{V}^t_{\mathcal{T}_i} = 1$)

2] $V^t_{\mathcal{T}_i} = 0$ due to falsification: "likelihood" $U^t_{\mathcal{T}_i} = 0$ (implying "verifiability" $T^t_{\mathcal{T}_i} > 0$):

d) ($V^t_{\mathcal{T}_i} = 0$) and $F^t_{\mathcal{T}_i} > 0$ and another theory has $V^t_{\mathcal{T}_{i \neq j}} > 0 \Longrightarrow \Pi^t_{\mathcal{T}_i} = 0 \Longrightarrow \mathcal{T}_i$ is false ($\underline{V}^t_{\mathcal{T}_i} = 0$) ...and if $F^t_{\mathcal{T}_i} >> 0 \Longrightarrow \mathcal{T}_i$ is very false

e) ($V^t_{\mathcal{T}_i} = 0$) and $F^t_{\mathcal{T}_i} > 0$ no other theory has $V^t_{\mathcal{T}_{i \neq j}} > 0 \Longrightarrow \Pi^t_{\mathcal{T}_i} = \frac{0}{0}$: $\mathcal{T}_i$ has indeterminate probability $\Longrightarrow \mathcal{T}_i$ is not a truth-bearer. But it's always conceivable a theory equal to what has been falsified, which would get $V^t_{\mathcal{T}_j} > 0 \Longrightarrow \Pi^t_{\mathcal{T}_i} = 0 \Longrightarrow \mathcal{T}_i$ is false ($\underline{V}^t_{\mathcal{T}_i} = 0$)

3] $V^t_{\mathcal{T}_i} = 0$ due to non-verifiability: "verifiability" $T^t_{\mathcal{T}_i} = 0$ (implying "likelihood" $U^t_{\mathcal{T}_i} = 1$):

f) $V^t_{\mathcal{T}_i} = 0$ and $F^t_{\mathcal{T}_i} = 0$ and another theory has $V^t_{\mathcal{T}_{i \neq j}} > 0 \Longrightarrow \Pi^t_{\mathcal{T}_i} = 0 \Longrightarrow \mathcal{T}_i$ is false ($\underline{V}^t_{\mathcal{T}_i} = 0$)

g) $V^t_{\mathcal{T}_i} = 0$ and $F^t_{\mathcal{T}_i} = 0$ and no other theory has $V^t_{\mathcal{T}_{i \neq j}} > 0 \Longrightarrow \Pi^t_{\mathcal{T}_i} = \frac{0}{0}$: $\mathcal{T}_i$ has indeterminate probability $\Longrightarrow \mathcal{T}_i$ is not a truth-bearer