Deceptive Credences (February, 2020 - forthcoming *ERGO*)

Ruobin Gong (Statistics, Rutgers University), Joseph B. Kadane (Statistics, CMU), Mark J. Schervish (Statistics, CMU), Teddy Seidenfeld (Philosophy and Statistics, CMU), and Rafael B. Stern (Statistics, Federal University of São Carlos)

Abstract

A familiar defense of Personalist or Subjective Bayesian theory is that, under a variety of sufficient conditions, asymptotically – with increasing shared evidence – *almost surely*, each non-extreme, countably additive Bayesian opinion, when updated by conditionalization, converges to certainty that is veridical about the truth/falsity of hypotheses of interest. Then, with probability 1 over possible evidential histories, personal probabilities track the truth. In this note we examine varieties of failures of these asymptotics. In an extreme case, conditional probabilities are *deceptive* when they converge to certainty for a false hypothesis.   We establish that proposals for so-called "modest" credences, offered by Elga (2016) and by Nielsen and Stewart (2019) in response to a concern about Bayesian *orgulity* raised by Belot (2013), instead support *deceptive* credences. We argue that deceptive credences are not *modest*, but for a reason different than Belot adduces.

1. *Introduction*.

In this note we continue an old discussion of some familiar results about the asymptotics of Bayesian updating (aka *conditionalization[1]*) using countably additive[2] credences. One such result (due to Doob (1953), with details reported in Section 2) asserts that, for each hypothesis of interest H, with the exception of a probability 0 "null" set of data sequences, the Bayesian agent's posterior probabilities converge to the truth value of H. Almost surely, the posterior

---

[1] To model changes in personal probability when learning evidence *e*, Bayesian *conditionalization* requires using the current conditional probability function P($\cdot$ | $\cdot$, *e*) as the updated conditional probability P′($\cdot$ | $\cdot$) upon learning evidence *e*.

[2] We use the languages of events to express these conditions. Let P($\cdot$) be a probability function. Let $E_1, \ldots, E_k$ be k-many pairwise disjoint events and E their union: $E_i \cap E_j = \varnothing$ if $i \neq j$, and E = $\cup_i E_i$.   Finite additivity requires: P(E) $= \sum_{i=1}^{k} P(E_i)$.   Let $E_1, \ldots, E_k, \ldots$ be countably many pairwise disjoint events and E their union: $E_i \cap E_j = \varnothing$ if $i \neq j$, and E = $\cup_i E_i$.   Countable additivity requires: P(E) $= \sum_{i=1}^{\infty} P(E_i)$.

credences converge to the value 1 if *H* is true, and to 0 if *H* is false. So, with probability 1, this Bayesian agent's asymptotic conditional credences are veridical: they track the truth of each hypothesis under investigation. This feature of Bayesian learning is often alluded to in a justification of Bayesian methodology, e.g., Lindley (2006, chapter 11) and Savage (1972, section 3.6): Bayesian learning affords sound asymptotics for scientific inference.

In Section 3, we explore the asymptotic behavior of conditional probabilities when these desirable asymptotics fail and credences are not veridical. We identify and illustrate five varieties of such failures, in increasing severity. An extreme variety occurs when conditional probabilities approach certainty for a false hypothesis. We call these extreme cases episodes of *deceptive* credences, as the agent is not able to discriminate between becoming certain of a truth and becoming certain of a falsehood.[3] *Result 1* establishes a sufficient condition for credences to be deceptive. *In Appendix A*, we discuss four other, less extreme varieties when conditional probabilities are not veridical.

In Section 4 we apply our findings to a recent exchange prompted by Belot's (2013) charge that familiar results about the asymptotics of Bayesian updating display *orgulity*: an *epistemic immodesty* about the power of Bayesian reasoning. In rebuttal, Elga (2016) argues that *orgulity* is avoided with some merely finitely additive credences for which the conclusion of Doob's theorem is false. Nielsen and Stewart (2019) offer a synthesis of these two perspectives where some finitely additive credences display what they call (understood as a technical term) *reasonable modesty*, which avoids the specifics of Belot's objection. Our analysis in Section 4 shows that these applications of finite additivity support *deceptive* credences . We argue that it is at least problematic to call deceptive credences "modest" in the ordinary sense of the word 'modest' when deception has positive probability.

---

[3] Deceptive credence is a worse situation for empiricists than what James (1896, Section 10) notes, where he famously writes,

> But if we are empiricists [pragmatists], if we believe that no bell in us tolls to let us know for certain when truth is in our grasp, then it seems a piece of idle fantasticality to preach so solemnly our duty of waiting for the bell.

It is not merely that the investigator fails to know when, e.g., her/his future credences for an hypothesis remain forever within epsilon of the value 1. With deceptive credences, the agent conflates asymptotic certainty of true statements with asymptotic certainty of false statements. The two cases become indistinguishable!

2. [Doob's (1953)]{.underline} strong law for asymptotic Bayesian certainty.

For ease of exposition, we use a continuing example throughout this note. Consider a Borel space of possible events based on the set of denumerable sequences of binary outcomes from flips of a coin of unknown bias using a mechanism of unknown dynamics. The sample space consists of denumerable sequences of 0s (tails) and 1s (heads). The nested data available to the Bayesian investigator are the growing initial histories of length n, h_n, arising from one denumerable sequence of flips, which corresponds to the unknown state. The class of hypotheses of interest are the elements of the Borel space generated by such histories.

For example, an hypothesis of interest H might be that, with the exception of some finite initial history, the observed relative frequency of 1's remains greater than 0.5, regardless whether or not there is a well defined limit of relative frequency for heads. Doob's result, which we review below, asserts that for the Bayesian agent with countably additive credences P over this Borel space, with the exception of a P-null set of possible sequences, her/his conditional probabilities, $P(H \mid h_n)$ converge to the truth value of H.

Consider the following, strong-law (countably additive) version of the Bayesian asymptotic approach to certainty, which applies to the continuing example of denumerable sequences of 0s and 1s.[4] The assumptions for the result that we highlight below involve the *measurable space*, the *hypothesis* of interest, and the *learning rule*.

The measurable space $\langle \mathcal{X}, \mathcal{B} \rangle$. Let X_i ($i = 1, \ldots$) be a denumerable sequence of sets, each equipped with an associated, atomic σ-field $\mathcal{B}_i$, where if $x_i \in X_i$ then $\{x_i\} \in \mathcal{B}_i$. That is, the elements of X_i are the atoms of $\mathcal{B}_i$. X_i is the state-space and $\mathcal{B}_i$ is the set of the measurable events for the ith experiment. Form the infinite Cartesian product $\mathcal{X} = X_1 \times X_2 \times \ldots$ of all sequences $x = (x_1, x_2, \ldots)$, where $x_i \in X_i$. The σ-field $\mathcal{B}$ is generated by the measurable rectangles from $\mathcal{X}$: the sets of the form $A = A_1 \times A_2 \times \ldots$ where $A_i \in \mathcal{B}_i$ and $A_i = X_i$ for all but finitely many values of i. $\mathcal{B}$

---

[4] See, also, Theorem 2, Section IV of Schervish and Seidenfeld (1990).

is the smallest σ-field containing each of the individual $\mathcal{B}_i$. As $\{x_i\} \in \mathcal{B}_i$ for each $x_i \in X_i$, also $\mathcal{B}$ is atomic with atoms the sequences x.

Each <u>hypothesis of interest</u> H is an element of $\mathcal{B}$. That is, in what follows, the result about asymptotic certainty applies to an hypothesis H provided that it is "identifiable" with respect to the σ-field, $\mathcal{B}$, generated by finite sequences of observations.[5] These finite sequences constitute the observed data.

We are concerned, in particular, with tracking the nested histories of the initial n experimental outcomes: $\qquad\qquad h_n = (x_1, \ldots, x_n)$, for $n = 1, 2, \ldots$

That is, for $x = (x_1, x_2, \ldots) \in \mathcal{X}$, let $h_n(x) = (x_1, \ldots, x_n)$ be the first n-terms of x.

The <u>probability assumptions</u>. Let P be a countably additive probability over the measurable space $\langle \mathcal{X}, \mathcal{B}\rangle$, and assume there exist well-defined conditional probability distributions over hypotheses $H \in \mathcal{B}$, given the histories h_n: $P(H \mid h_n)$, $n = 1, \ldots$.

The <u>learning rule for the Bayesian agent</u>: Consider an agent whose initial ("prior") joint credences are represented by the measure space $\langle \mathcal{X}, \mathcal{B}, P \rangle$. Let P^n be this agent's ("posterior") credences over $\langle \mathcal{X}, \mathcal{B}\rangle$ having learned the history h_n.

Bayes' Rule for updating credences requires that $\qquad P^n(H) = P(H \mid h_n)$.

The result in question, which is a substitution instance of Doob's (1953 T.7.4.1), is as follows: For $H \in \mathcal{B}$, let $\mathbf{I}_H\colon \mathcal{X} \to \{0,1\}$ be the indicator for H. $\mathbf{I}_H(x) = 1$ if $x \in H$ and $\mathbf{I}_H(x) = 0$ if $x \notin H$. The indicator function for H identifies the truth value of H.

- *Asymptotic Bayesian Certainty*: For each $H \in \mathcal{B}$,

$$P\{x\colon \lim_{n \to \infty} P^n(H) = \mathbf{I}_H(x)\} = 1.$$

---

[5] See Schervish and Seidenfeld (1990), Examples 4a and 4b for illustrations where H is not an element of $\mathcal{B}$ and where the asymptotic certainty result fails.

In words, subject to the conditions above, the agent's credences satisfy asymptotic certainty about the truth value of the hypothesis H. For each measurable hypothesis H, and with respect to a set S_H of infinite sequences x that has "prior" probability 1, for each x in S_H her/his sequence of "posterior" opinions about H, $P(H|\ h_n(x))$, converges to probability 1 or 0, respectively, about the truth or falsity of H.

To summarize: For each x in S_H, as $n \to \infty$, the sequence of conditional probabilities, $P(H|\ h_n(x))$, asymptotically correctly identifies the truth of H or of H^c by converging to 1 for the true hypothesis in this pair. In this sense, asymptotically, the Bayesian agent learns whether H or H^c obtains.

   Definition: Call such states <u>*veridical*</u> about the truth of H versus H^c.[6]

In other words, the *non-veridical* states constitute the *failure set* for Doob's result.

3. *Veridical* versus *Deceptive* states and their associated credences.

Next, we examine details of conditional probabilities given elements of the failure set, even when the agent's credences are countably additive and the other assumptions in Doob's result obtain. Specifically, consider the countably additive Bayesian agent's conditional probabilities, $P(H|h_n)$, in sequences of histories that are generated by points x in the *failure set*, S_H^c – the complement to the distinguished set of veridical states. It is important, we think, to distinguish different varieties of non-veridical states within the failure set.

At the opposite pole from the veridical states, the states in S_H – states whose conditional probabilities converge to the truth about H – are states whose histories create conditional probabilities that converge to certainty about the *false* hypothesis in the pair $\{H, H^c\}$.

   Define x as a <u>*deceptive*</u> state for hypothesis H if $P(H|h_n(x))$ converges to $1 - \mathbf{I}_H(x)$.

For deceptive states, the agent's sequence of posterior probabilities also creates asymptotic certainty. This sense of certainty is introspectively indistinguishable to the investigator from the

---

[6] For ease of exposition, where the context makes evident the hypothesis *H* in question, we refer to states as *veridical* or *deceptive* simpliciter.

asymptotic certainty created by veridical states, where asymptotic certainty identifies the truth. Thus, to the extent that veridical states provide a defense of Bayesian learning – the observed histories $h_n(x)$ move the agent's subjective "prior" for H towards certainty in the truth value of H – deceptive states move the agent's subjective credences towards certainty for a falsehood. Thus, for the very reasons that states in S_H underwrite a Bayesian account of Bayesian learning of H, deceptive states frustrate such a claim about H. Then, Doob's result serves a Bayesian's need provided that the Bayesian agent is satisfied that, with probability 1, the actual state is *veridical* rather than *deceptive* with respect to the hypothesis of interest.

When the failure set for an hypothesis H is deceptive, then the investigator's credences about H converge to 0 or to 1 for all possible data sequences. But this convergence is logically independent of the truth of H since the investigator is unable to distinguish veridical from non-veridical data histories.

Less problematic than being deceptive, but nonetheless still challenging for a Bayesian account of objectivity, is a non-deceptive state x where for each $\varepsilon > 0$, infinitely often

$$|P(H|h_n(\mathrm{x})) - \mathbf{I}_H(x)| > 1\text{-}\varepsilon.$$

Then, with respect to hypothesis H, infinitely often x induces non-veridical conditional probabilities that mimic those from a deceptive state.

    Definition: Call such a state *intermittently deceptive* for hypothesis H.

Consider a non-veridical state where, for each $\varepsilon > 0$, infinitely often $|P(H|h_n(\mathrm{x})) - \mathbf{I}_H(x)| < \varepsilon$.

    Definition: Call such a state *intermittently veridical* for hypothesis H.

Within the failure set for an hypothesis, the following partition of non-veridical states appears to us as increasingly problematic for a defense of Bayesian methodology, in the sense that seeks asymptotic credal certainty about the truth value of the hypothesis driven by Bayesian learning. In this list, we prioritize avoiding deception over obtaining veridicality.[7]:

---

[7] We note in passing that the categories may be further refined by considering sojourn times for events that are required to occur infinitely often. Also, the categories may be expanded to include, δ-*veridical* and δ-*deceptive*, where for some $\delta > 0$, conditional probabilities, $P(H|h_n(x))$, accumulate (respectively) to within δ of $\mathbf{I}_H(x)$ and to within δ of $1 - \mathbf{I}_H(x)$. We do not consider these variations here.

(A) states that are intermittently veridical but not intermittently deceptive;

(B) states that are neither intermittently veridical nor intermittently deceptive;

(C) states that are both intermittently veridical and intermittently deceptive[8];

(D) states that are intermittently deceptive but not intermittently veridical;

(E) states that are deceptive.

We find it helpful to illustrate these categories within the continuing example of sequences of binary outcomes. Consider the set of denumerable, binary sequences: $\mathcal{X} = \{x: \mathbf{N}^+ \rightarrow \{0,1\} \}$. That is, in terms of the structural assumptions in Doob's result, $X_i = \{0,1\}$; each $\mathcal{B}_i$ is the 4-element algebra $\{\varnothing, \{0\}, \{1\}, \{0,1\} \}$, for $i = 1, 2, \ldots$; and the inclusive σ-field $\mathcal{B}$ is the Borel σ-algebra generated by the product of the $\mathcal{B}_i$.

First, if H is defined by finitely many coordinates of x (a finite dimensional rectangular event) then $\mathbf{P}^n(H)$ converges to the indicator function for H, $\mathbf{I}_H$, after only finitely many observations. Then $S_H = \mathcal{X}$ and all states are veridical. That is, there is no sequence where the conditional probabilities $\mathbf{P}^n(H)$ fail to converge to $\mathbf{I}_H$. Moreover, this situation obtains regardless whether P is countably or merely finitely additive, provided solely that $P(E \mid h_n)$ is a conditional probability that satisfies the following *propriety* condition: $P(B|A) = 1$ whenever $\varnothing \neq A \subseteq B$.

Next, consider an hypothesis that is logically independent of each finite dimensional rectangular event, an hypothesis that is an element of the *tail* σ-sub-field of $\mathcal{B}$. For instance, note that each sequence x has a well-defined *lim inf* and *lim sup* of the relative frequency for the digit 1. Let $L(x) = l$ be this *lim inf* and $U(x) = u$ be this *lim sup*. Then the set of ordered pairs $\{<l,u>: 0 \leq l \leq u \leq 1\}$ partitions $\mathcal{X}$ into $\mathcal{B}$-measurable events. Each ordered pair $<l,u>$ corresponds to a $\mathcal{B}$-

---

[8] Our understanding is that case (C) satisfies the conditions for what Belot (2013) calls a "flummoxed" credence. Weatherson (2015) discusses varieties of "open minded" credences, including those that are "flummoxed," in connection with Imprecise Probabilities. Here, we focus on failures of veridicality for coherent, precise credences.

measurable subset of $\mathcal{X}$ with cardinality of the continuum. *Figure 1*, below, graphs these points in the isosceles right triangle with corners <0,0>, <1,1> and <0,1>.

FIGURE 1

Isosceles right triangle with corners <0,0>, <1,1> and <0,1>
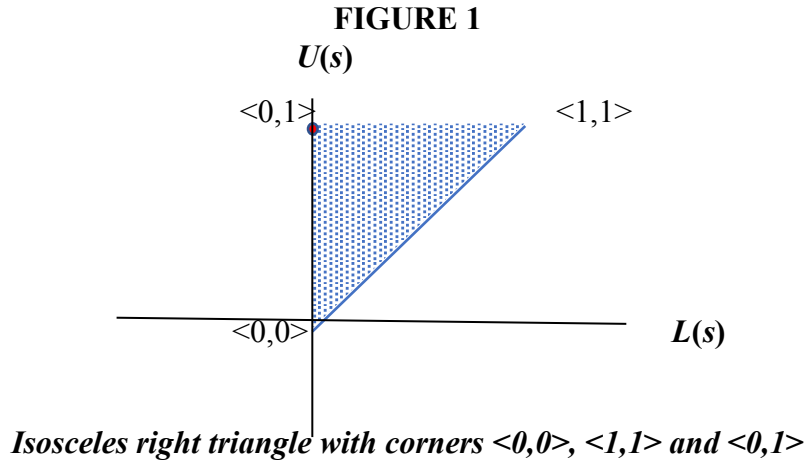
Let H be the subset of $\mathcal{X}$ of sequences with a well-defined limit of relative frequency for the digit 1. In Figure 1, H corresponds to the set of ordered pairs <l,u> with $l = u$, the (blue) line of points along the main diagonal.

For a countably additive personal probability that satisfies de Finetti's (1937) condition of exchangeability, this subset H of $\mathcal{X}$ has personal "prior" probability 1, $P(H) = 1$. Also, assume for convenience that this probability P is not extreme within the class of exchangeable probabilities: $0 < P(\{1\}) < 1$. Then for each sequence x in $\mathcal{X}$, $P(h_n(x)) > 0$, and trivially, also $P(H|h_n(x)) = 1$. For the result on asymptotic Bayesian certainty, then $S_H = H$. However, on the complementary set, for $x \in S_H^c$ the conditional probabilities satisfy: $P(H \mid h_n(x)) = 1$; hence, each $x \in S_H^c$ is deceptive: category (E). Moreover, under these conditions, when a state is not veridical then it is *deceptive*: the posterior probability converges to $1 - \mathbf{I}_H(x)$.

Definition: Call a failure set S_H^c *deceptive* if each state in the failure set is deceptive for H.

        Also, in this case we say that the associated credence is *deceptive* for H.

We summarize this elementary finding as follows:

> ***Result 1*** Suppose that the credence function treats each possible initial history h_n as not
> "null": $P(h_n(x)) > 0$. Then for each hypothesis H ($\neq \Omega$) for which $P(H) = 1$, the failure-set
> for H is not empty and deceptive.

Moreover, if the space is uncountable, so that there is an uncountable partition of the space each
of whose elements is an uncountable set, as depicted in Figure 1, then we have the following as
well:

> ***Corollary***: For each finitely additive probability P on a space of denumerable sequences
> of (logically independent) random variables, where each initial history h_n is not "null,"
> there exists an hypothesis H, with $P(H) = 1$, whose failure set S_H^c is an uncountable set,
> and that failure set is deceptive.

The non-veridical states, $x \in S_H^c$, can populate each of the other four categories, (A) – (D). We
discuss these in Appendix A.

4. *Reasonably modest* but *deceptive* failure sets.

Next, we apply these finding to a recent debate about what Belot (2013) alleges is mandatory
Bayesian *orgulity*. We understand Belot's meaning as follows. For a Bayesian agent who
satisfies, e.g., the conditions for Doob's result, the set of samples where the desired asymptotic
certainty <u>fails</u> for an hypothesis H (the so-called "*failure set*" for H) has probability 0.
Nonetheless, this failure set may be a "large" or "typical" event when considered from a
topological perspective. Specifically, the failure set may be comeager with respect to a
privileged product topology for the measurable space of data sequences. As we understand
Belot's criticism, such a Bayesian suffers *orgulity* because she/he is obliged by the mathematics
of Bayesian learning to assign probability 0 to the possible evidence where the desired
asymptotic result fails, even when this failure set is comeager.

In a (2016) reply to Belot's analysis, A. Elga focuses on the premise of countable additivity in
Doob's result. Countable additivity is not required in either Savage's (1972) nor de Finetti's
(1974) theories of Bayesian coherence. Elga gives an example of a merely finitely additive (and
not countably additive) probability over denumerable binary sequences and a particular
hypothesis H where with positive probability (in fact, with probability 1) the investigator's

posterior probability fails to converge to the indicator function for *H*. So, not all finitely additive coherent Bayesians display *orgulity*.

M. Nielsen and R. Stewart (2019) extend the debate by explicating what they understand to be Belot's rival account of *reasonable modesty* of Bayesian conditional probabilities. They offer a reconciliation of Elga's rebuttal and Belot's topological perspective.   For Nielsen and Stewart, a credence function is *modest* for an hypothesis *H* provided that it gives (unconditional) positive probability to the *failure set* for the convergence of posterior probabilities to the indicator function for *H*. By this account, each credence in the class of countably additive credences is *immodest* over all hypotheses that are the subject of the asymptotic convergence result but have non-empty failure sets. Since requiring *modesty* for all such hypotheses is too strong of a condition even for (merely) finitely additive credences – as per the Corollary to Result 1, above – Nielsen and Stewart propose a standard of *reasonable modesty*. This condition requires *modesty* solely for failure sets that are typical in the topological sense, for some privileged topology.

With their Propositions 1 and 2, Nielsen and Stewart point out that there exist a class of merely finitely additive credences (with cardinality of the continuum) such that each credence function in this class assigns unconditional positive probability (even probability 1) to each comeager set. Then, such a credence displays *reasonable modesty* for each failure set that is "typical."

Below, we show that the *reasonably modest* credences that Nielsen and Stewart point to with their Proposition 1, nonetheless, mandate deceptive failure sets for specific hypotheses. And as we explain (in Appendix B), Nielsen and Stewart's Proposition 2 provide *reasonably modest* credences in their technical sense at the price of making it impossible to learn about hypotheses that concern unobserved parameters, in all familiar statistical models.

First we argue that this sense of "modesty" is mistaken when deception is not a null event, regardless whether the *modesty* is *reasonable* or not. When the investigator's credences are merely finitely additive, with respect to a particular hypothesis the failure set for Doob's result

may have positive prior probability, as is well known.[9] In such cases, the investigator's credences are called *modest* according to Nielsen and Stewart. Suppose, further, that such a modest credence also has a deceptive failure set. Then, each state is either veridical or deceptive. But the investigator <u>behaves</u> just as though asymptotic certainty tracks the truth. That is, the fact that the set of deceptive states (for a particular hypothesis) has positive probability – $P(S_H^c) > 0$ rather than $P(S_H^c) = 0$ – the fact that the investigator's credence is "modest," is irrelevant to the investigator's decision making. Here is why.

Let H be an hypothesis, and suppose that each state is either veridical for H or deceptive for H. Then, for each state x, the sequence $\{P(H \mid h_n(x)): n = 1,2,\ldots\}$ converges to 1 *if and only if* either x is veridical and in H, or if x is deceptive and in H^c. And $\{P(H \mid h_n(x)): n = 1,2,\ldots\}$ converges to 0 *if and only if x* is veridical and in H^c, or if x is deceptive and in H. Hence, the investigator becomes asymptotically certain about the truth of H no matter what data are observed. This analysis holds regardless of what prior probability the investigator assigns to H and regardless how probable is the failure set. The *modesty* of P for H, namely that $P(S_H^c) > 0$, is irrelevant to this conclusion. And so too, it is irrelevant to this conclusion whether the *modesty* of P for H is *reasonable* or not. It is irrelevant whether S_H^c is a comeager set or not.

To put this analysis in behavioral terms, suppose the Bayesian investigator faces a sequence of decisions. These decisions might be practical, with cardinal utilities that reflect economic or legal, or ethical consequences. Or, these decisions might be cognitive with epistemically motivated utilities, e.g., for desiring true hypotheses over false ones, or for desiring more informative over less informative hypotheses. Or, these might form a mixed sequence of decisions, with some practical and some cognitive. Suppose each decision in this sequence rides on the probability for one specific hypothesis H and, regarding the corresponding sequence of Bayesian conditional probabilities for H that parallel these decisions, the investigator's credence is deceptive for H. Then, asymptotically, the investigator's sequence of decisions will be determined by the asymptotic certainty – the conditional credence for H of 0 or 1 – that surely results, no matter which sequence of observations obtains. But if also the investigator has

---

[9] Moreover, when credences are merely finitely additive, the investigator may design an experiment to insure deceptive Bayesian reasoning. For discussion see Kadane et al. (1996).

a positive unconditional probability for deception, this "modesty" plays no role in her/his sequence of decisions. The "modesty" reported by her/his unconditional probability of deception, $P(S_H^c) > 0$, be it a large or a small positive probability, is irrelevant to the sequence of decisions that she/he makes. When a failure set is both deceptive and non-null, the Bayesian investigator ignores this in her/his decision making, treating all certainties alike. Just as if $P(S_H^c)$ = 0. We do not agree, then, that the investigator's credences are *modest* for hypothesis H when the failure set is deceptive and $P(S_H^c) > 0$.

One example in which the conditions of this analysis hold was given by Elga (2016) and is an instance of our continuing example about binary sequences. In Elga's example, H is the hypothesis that the binary sequence satisfies $L(x) = U(x) = .9$. In his example the failure set S_H^c is deceptive with probability 1, i.e., $P\{x: x$ is deceptive for $H\} = 1$.[10]

A large class of examples of this kind arise by using Proposition 1 of Nielsen and Stewart. Here is how Proposition 1 applies to the continuing example of the Borel space, $\mathscr{B}$, of binary sequences on $\{0,1\}$. Let P_1 be a non-extreme, exchangeable countably additive probability. That is, in addition to being an exchangeable probability, for each finite initial history, i.e., for each of the 2^n possible sequences h_n, and for each $n = 1, 2, \ldots$, then $P_1(h_n) > 0$. By Doob's result, P_1 is not *modest* (in Nielsen and Stewart's sense) because, for each hypothesis H its failure set is P_1-null, $P_1(S_H^c) = 0$. Let P_2 be a finitely additive, 0-1 ("ultrafilter") probability with the property that if E is a comeager set in $\mathscr{B}$, then $P_2(E) = 1$.[11] Fix $0 < y < 1$ and define $P = yP_1 + (1-y)P_2$, the $y:(1-y)$ mixture of these two probabilities.

---

[10] By contrast, in Cisewski et al's (2018) version of Elga's example, for the same hypothesis H, the failure set, S_H^c = **X**, is the whole space; whereas, for each x and for each n, $P(H|h_n(x)) = \frac{1}{2} = P(H)$. Then the failure set generates solely indecisive conditional credences: each state is neither intermittently veridical nor intermittently deceptive – category (B).

[11] Existence of such 0-1 finitely additive probabilities is a non-constructive consequence (using the Axiom of Choice) that the comeager sets form a filter: They have the finite intersection property and are closed under supersets.

Nielsen and Stewart's Proposition 1 establishes that P is *reasonably modest*, since for each hypothesis H, if the failure set S_H^c is comeager, then $P(S_H^c) > 0$. However, as we show next, Proposition 1 creates *reasonably modest* credences that, in the Continuing Example, have failure sets for specific hypotheses that have positive probability, are comeager, and are deceptive.

> **Result 2** In the continuing example, let H be the hypothesis that the binary sequence belongs to the set of maximally chaotic relative frequencies, corresponding to the (red) point <0,1> in Figure 1. This is the set of sequences with *lim inf* (rel freq "1") = 0 and *lim sup* (rel freq "1") = 1. Then the failure set for H under P, S_H^c, has positive probability, $P(S_H^c) = (1-y) > 0$, is comeager, and is deceptive.

Proof: Because both $P_1(H) = 0$ and for each history h_n, $P_1(h_n) > 0$, then $P_1(H \mid h_n) = 0$. Under P_2 there is a distinguished binary sequence x_{P_2} in the following sense. The finite initial histories form a binary branching tree: for each n there are 2^n distinct histories h_n. Because P_2 is an "ultrafilter" distribution, then for each n and for each possible finite initial history h_n of length n, $P_2(h_n) = 0$ or $P_2(h_n) = 1$. So, there is one and only one sequence x_{P_2} where, for each n, $P_2(h_n(x_{P_2})) = 1$.[12] That is, for each sequence $x' \neq x_{P_2}$ there exists an m such that for all $n > m$, $P_2(h_n(x')) = 0$. Thus, for each $x' \neq x_{P_2}$ there exists an m such that for all $n > m$,

$$P(H \mid h_n(x')) = P_1(H \mid h_n(x')) = 0.\text{[13]}$$

Specifically, the failure set S_H^c is either the set $H - \{x_{P_2}\}$ (if the sequence x_{P_2} belongs to H), or it is the set $H \cup \{x_{P_2}\}$ (if the sequence x_{P_2} belongs to H^c). In either case, the failure set S_H^c is deceptive for H. According to Cisewski et al. (2018) H is a comeager set. Evidently then, S_H^c is a comeager set where $P(S_H^c) = (1-y)P_2(S_H^c) = (1-y)P_2(H) = (1-y) > 0$.[14] QED

---

[12] Note well that P_2 is merely finitely additive as $P_2(x_{P_2}) = 0$, since each unit set $\{x\}$, each denumerable sequence x, is a meager set.

[13] More generally, if $x' \neq x_{P_2}$ the agent's conditional probabilities become and stay *immodest*, as they become the sequence of countably additive conditional probability function, $P_1(\cdot \mid h_n(x'))$. So, though P is *modest*, with P-probability 1 its conditional credences become and stay *immodest*.

[14] Similarly, Result 2 applies to each hypothesis H of a comeager set whose complement includes the support of the countably additive, immodest probability P_1.

We emphasize that certainty with deception is indistinguishable from certainty that is veridical. In the context of Result 2, the investigator can tell when the observed history h_n differs from the history that would be observed in the one distinguished sequence, $h_n(x_{P_2})$. But that recognition provides no basis for altering the certainty, $P(H \mid h_n) = 0$, that results once the observed history departs from the distinguished one, once $h_n \neq h_n(x_{P_2})$. Regardless the magnitude of the (unconditional) probability of deception, $P(S_H^c)$, the investigator cannot identify when certainty is deceptive rather than when it is veridical. Her/his conditional credence function, $P(\cdot \mid h_n)$, already takes into account the total evidence available. Certainty is certainty, full stop.

We have argued above that a credence P is not epistemically modest where there is an hypothesis H that has a deceptive failure set S_H^c that is not P-null. Then, in the continuing example, each probability P created according to Proposition 1 fails this test of epistemic modesty.

In Summary, it is our view that having a positive probability over non-veridical states is not sufficient for creating an epistemically modest credence because categories (D) or (E) may have positive prior probability as well. Indeed, in the continuing example, each probability P created according to Proposition 1 fails this test of epistemic modesty.

5. We summarize the principal conclusion of this note:
- When the failure set for an hypothesis is deceptive and not null, that is in conflict with an attitude of epistemic modesty about learning that hypothesis.

Regarding the asymptotics of Bayesian certainties, e.g. Doob's result, neither of Nielsen and Stewart's concepts of *modesty*, nor *reasonable modesty* distinguishes deceptive from other varieties of failure sets. According to Result 2, in the Continuing Example each credence P that satisfies Nielsen and Stewart's Proposition 1 admits an hypothesis whose failure set is P-non-null, comeager, and deceptive.

References

Belot, Gordon (2013) *Bayesian Orgulity*. *Phil. Sci.* **80**: 483-503.

Cisewski, Jessica, J. B. Kadane, M. J. Schervish, T. Seidenfeld, and R. B. Stern (2018) *Standards for Modest Bayesian Credences*. *Phil. Sci.* **85**: 53-78.

de Finetti, Bruno (1937) *Foresight: Its Logical Laws, Its Subjective Sources*. In Henry E. Kyburg, Jr. and Howard E. Smokler (eds) *Studies in Subjective Probability* (1964). John Wiley.

de Finetti, Bruno (1974) *Theory of Probability*, Vol. 1. John Wiley.

Doob, Joseph L. (1953) *Stochastic Processes*. Jphn Wiley.

Elga, Adam (2016) *Bayesian Humility*. *Phil. Sci.* **83**: 305-323.

James, William (1896) *The Will to Believe*, *The New World*: **5**: 327-247. Reprinted in his (1962) *Essays on Faith and Morals*. The World Publishing Co.

Kadane, Joseph B., M. J. Schervish, and T. Seidenfeld (1996) *Reasoning to a Foregone Conclusion*. *JASA* **91**: 1228-1236.

Lindley, Dennis V. (2006) *Understanding Uncertainty*. John Wiley.

Nielsen, Michael, and R. Stewart (*2019*) *Obligation, Permission, and Bayesian Orgulity*. *Ergo* **6**: 58-70.

Savage, Leonard J. (1972) *The Foundations of Statistics* (2nd revised edition). Dover Press.

Schervish, Mark J., and T. Seidenfeld (1990) *An Approach to Certainty and Consensus with Increasing Evidence*. *J. Stat. Planning and Inference* **25**: 401-414.

Weatherson, Brian (2015) *For Bayesians, Rational Modesty Requires Imprecision*. *Ergo* **2**: 529-545.

Appendix A

Here, we discuss and illustrate categories (A) – (D) of failure sets using the continuing example. Restrict the exchangeable "prior" probability P so that, in terms of de Finetti's *Representation Theorem*, the "mixing prior" for the Bernoulli parameter is smooth, e.g. let it be the uniform $U[0, 1]$. Choose $0 < c < d < 1$ and consider the hypothesis $H = \{x: c \leq L(x) \leq U(x) \leq d\}$. So, with the "uniform" prior, $P(H) = d\text{-}c$; so, $1 > P(H) > 0$.

The set of veridical states for this credence and hypothesis includes each sequence where,

  either $c < L(x) \leq U(x) < d$ – in which case H obtains and $lim_{n \to \infty} P(H \mid h_n) = 1$,

  or, either $U(x) < c$ or $L(x) > d$ – in which case H^c obtains and $lim_{n \to \infty} P(H \mid h_n) = 0$.[15]

The non-veridical states (the failure set) S_H^c, the set of sequences where $P(H \mid h_n(x))$ does not converge to the indicator $\mathbf{I}_H(x)$, include states x such that $L(x) < c < U(x)$ or $L(x) < d < U(x)$. For such a state x, $P(H|h_n(x))$ fails to converge and

$$lim \; inf \; P(H|h_n(x)) = 0 \text{ and } lim \; sup \; P(H|h_n(x)) = 1.$$

Then x is both intermittently veridical and intermittently deceptive for H – category (C).

In order to illustrate the other three categories of non-veridical states, (A), (B), and (D), the following adaptation of the previous construction suffices. Depending upon which category is to be displayed, consider a state x such that the likelihood ratio

$$P(h_n(x)|H) / P(h_n(x)|H^c)$$

oscillates with suitably chosen bounds, in order to have the sequence of posterior odds,

$$P^n(H) / P^n(H^c)$$

oscillate to fit the category. This method succeeds because, as is familiar, the posterior odds equals the likelihood ratio times the prior odds:

$$P^n(H) / P^n(H^c) = [P(h_n(x)|H) / P(h_n(x)|H^c)] \times [P(H) / P(H^c)].$$

---

[15] When either $c = L(x)$ and $U(x) < d$, or $c < L(x)$ and $U(x) = d$, or $c = L(x)$ and $U(x) = d$, then the behavior of $lim_{n \to \infty} P(H \mid h_n)$ is not determined. This issue is relevant to the illustration of case (A), with clause (*ii*), below.

We illustrate category (A) using the same hypothesis $H = \{x: c \le L(x) \le U(x) \le d\}$ and credence as above. For a non-veridical state in category (A), consider a sequence x such that both:

(i) $c < U(x) < d$. Then x is intermittently veridical as, infinitely often, the relative frequency of '1' falls strictly between c and d, and

(ii) $L(x) = c$ but there exists $0 < \rho < \infty$, where for only finitely many values of n,

$$P(h_n(x)|H) / P(h_n(x)|H^c)] \ < \ \rho - \text{so that } x \text{ is not intermittently deceptive,}$$

and infinitely often $P(h_n(x)|H) / P(h_n(x)|H^c)] = \rho$ – so that x is not veridical.

Appendix B

In this appendix we consider Nielsen and Stewart's Proposition 2, and related approaches for creating a *reasonably modest* credence, P'. We adapt Proposition 2 to the continuing example of the Borel space of denumerable binary sequences. Consider a finitely additive probability P' on the space of binary sequences in accord with Nielsen and Stewart's *Proposition 2*, where

(i) $P'(h_n) > 0$ for each possible finite initial history;

and (ii) $P'(E) = 1$, whenever E comeager.

Nielsen and Stewart's Proposition 2 asserts that, however P' is defined on the field of finite initial histories, which space we denote by $\mathcal{A}$, $\mathcal{A} \subset \mathcal{B}$, then P' may be extended to a finitely additive probability that is extreme with respect to the field of comeager and meager sets in $\mathcal{B}$.

For example, if P_1 is a countably additive probability on $\mathcal{B}$, then P' might agree with P_1 on $\mathcal{A}$, while $P'(E) = 1$ if E is a comeager set. Then, P' is *reasonably modest* in the technical sense used by Nielsen and Stewart since, whenever a failure set S_H^c is comeager, $P'(S_H^c) = 1$.

We do not know whether the conclusion of *Result 2* extends also to the *reasonably modest* credences P' created according to the technique of Proposition 2. For instance, we do not know, for a general P', when an hypothesis H has a deceptive failure set S_H^c with $P'(S_H^c) > 0$. Evidently, we are unwilling to grant that a credence satisfying Proposition 2 is epistemically modest about learning an hypothesis H merely because $P'(S_H^c) > 0$ whenever S_H^c is a comeager set.

However, there is a second issue that tells against the technique of Proposition 2 for creating *reasonable modesty*. In Proposition 1, probability values from the *immodest* countably additive credence P_1 for events in the tail field of $\mathcal{B}$ are relevant to the values that the *reasonably modest* credence P gives these events. And, as P_1 is countably additive, the P_1 probability values for tail events are approximated by P_1 values in $\mathcal{A}$. In short, under the method used in Proposition 1, P_1 probability values for events in $\mathcal{A}$ constrain the *reasonably modest* values of $P(S_H^c)$. However, in Proposition 2 the P_1 values in $\mathcal{A}$ are not relevant to the P'-values for events in the tail field. In Proposition 2, the P' probability values are stipulated to be extreme for comeager sets, regardless how the P'-credences are assigned to the elements of the observable $\mathcal{A}$. The upshot is that with P' credences the investigator is incapable of learning about comeager sets based on Bayesian learning from finite initial histories.

With respect to the continuing example, Cisewski et al. (2018) establish that the set of sequences corresponding to the one point <0,1> in *Figure 1* is comeager. Thus, in order to assign a prior probability 1 to each comeager set, this agent is required to hold an extreme credence that the sequence has maximally chaotic relative frequencies: $P'\{x: x \in <0,1>\} = 1$.

As above, let the hypothesis of interest be $H = \{x: x \in <0,1>\}$: the hypothesis that the sequence has maximally chaotic relative frequencies. Then Result 1 obtains as $P'(H) = 1$ and $P'(H \mid h_n) = 1$ for each $n = 1, 2, \ldots$. No matter what the agent observes, her/his posterior credence about H remains extreme. With credence P', the failure set for H is the meager set (hence a P'-null set) of continuum many states corresponding to each point in Figure 1 other than the corner <0,1>. Each point in the failure set for H is deceptive: the failure set S_H^c is deceptive![16] On what basis do Nielsen and Stewart dismiss the deceptiveness of S_H^c as irrelevant to the question whether P' is an appropriate credence for investigating statistical properties of binary sequences? We speculate their answer is, solely, that the failure set S_H^c is meager.

---

[16] The Corollary to Result 1 establishes that the same phenomenon occurs when Nielsen and Stewart's Prop. 2 is generalized to include finitely additive credences that assign positive probability to each finite initial history and a positive (but not necessarily probability 1 credence) to each comeager set of sequences.

Propositions 1 and 2 do not exhaust the varieties of finitely additive probabilities that assign positive probability to each comeager set in $\mathcal{B}$. For instance, one may recombine the techniques from these two Propositions as follows.

Let P_1 be an (immodest) countably additive probability on $\mathcal{B}$ that assigns positive probability to each finite initial history. Let P_2 be a finitely additive probability defined on $\mathcal{B}$ obtained by the technique of Proposition 2, but where P_1 and P_2 agree on $\mathcal{A}$. So, $P_2(H) = 1$, for the hypothesis H that the sequence is maximally chaotic. Then, in the spirit of Proposition 1, define P_3 as a (non-trivial) convex combination of P_1 and P_2: let $0 < x < 1$ and define $P_3 = xP_1 + (1-x)P_2$. Then P_3 avoids the difficulty displayed by the probability P' of Proposition 2, discussed above, namely $P_3(H) = 1-x < 1$. There is no prior certainty under P_3 that the sequence is maximally chaotic.

But P_3 has its own difficulties. Here are two. The *Corollary* applies to P_3 with the hypothesis *H'*: that the sequence is either maximally chaotic or has a well defined limit of relative frequency. In *Figure 1*, H' corresponds to the sequences either in the set corresponding to the point <0,1> or in the set of points with well-defined limits of relative frequency, where $L(x) = U(x)$. The P_3 failure set for *H'* is uncountable and deceptive, though meager. Second, P_3 makes all observations irrelevant for learning about the hypothesis H: the sequence is maximally chaotic. This follows because $\qquad P_3(h_n \mid H) = P_2(h_n) = P_1(h_n) = P_3(h_n)$.

So, for each initial history, h_n

$$P_3(H \mid h_n) \ = \ P_3(h_n \mid H) \times P(H) \, / \, P_3(h_n) = P_3(H) = (1\text{-}x)$$

Independent of the history h_n.