addresses both issues of *certainty* and *consensus* among finitely many investigators over a finite partition of statistical hypotheses, assuming they share an increasing sequence of observations from random sampling.

Savage offers these findings as a partial defense against the accusation, voiced by frequentist statisticians of the time, that the theory of (Bayesian) personalist statistics is fraught with subjectivism and cannot serve the methodological needs of the Scientific community, where objectivity is required. The central theme in Savage's response is to understand 'objectivity' in terms of shared agreements about the truth, particularly, when the shared agreements arise from shared statistical evidence.

space is uncountable, instead require that they agree with each other about which events in this uncountably infinite space of observables have probability 0. They share in a family of *mutually absolutely continuous* probability distributions. If the agents' personal probabilities over these infinite spaces also are countably additive, then strong-law convergence theorems yield strengthened results about asymptotic consensus (see, e.g. Blackwell and Dubins, 1962) and also about asymptotic certainty for events defined in the space of sequences of increasing shared evidence. We discuss several of these results in Sectioneins

Though he does not explicitly formulate criteria for *immodesty*, based on the examples and analysis he offers, we understand Belot's primary requirements to be these two[6]:

- ***Topological***

humility. They promote excessive *apriorism* with respect to ordinary properties of limiting frequencies.

The Bayesian convergence-to-the-truth results that are the subject of Belot’s complaints are formulated as probability strong-laws that hold *almost surely* or *almost everywhere*. In order to make clear why we think Belot’s verdict is mistaken thinking these results about convergence to the truth are a liability for Bayesian theory, revisit the familiar instance of the strong law of large numbers, as reported in *fn*. 4.

Let <&, $\mathcal{B}$, P> be the countably additive measure space generated by all finite sequences of repeated, probabilistically independent [*iid*] flips of a “fair” coin. Let 1 denote a “Heads” outcome and 0 a “Tails” outcome for each flip. Then a point $\boldsymbol{x}$ of & is a denumerable sequence of 0s and 1s, $\boldsymbol{x} = <\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots >$, with each $\boldsymbol{x}_n$ {0, 1} for $n = 1, 2, \ldots$. Let $X_n(\boldsymbol{x}) = \boldsymbol{x}_n$ designate the random variable corresponding to the outcome of the n^{th} flip of the fair coin. $\mathcal{B}$ is the Borel (-algebra generated by *rectangular* events, those determined by specifying values for finitely many coordinates in &. P is the countably additive *iid* product *fair-coin* of the continuum.[8] When $2^{)}$ is equipped with the infinite product of the discrete

Unless a probability model P for a sequence of relative frequencies assigns probability 1 to the set of sequences of observed frequencies that oscillate maximally, then P assigns positive probability to a meager set of sequences, in violation of Condition #2. Evidently, the standard for epistemological *modesty* formalized in Topological Condition #2, which requires meager sets of relevant events be assigned probability 0, itself leads to probabilistic orgulity because it requires an unreasonable *a priori* opinion about how observed relative frequencies behave. Let P satisfy Condition #2. Given evidence of a P-non-null observation

frequency hypothesis in question: At each stage of her investigation, looking forward, she remains practically certain that her posterior probability will converge to the true limiting frequency hypothesis.

Second, the credal state P in Elga's example fails what we call Belot's Condition #1. P assigns probability 1 to a meager set of sequences of observations. Hence, though Elga argues that P is modest with respect to one

additive

since it is then following a $P^{1/10, 9/10}$ law.) Then, for each # > 0 there exists integer

$n_{\#}$

As required for Elga's construction, this finitely additive probability P behaves as $P^{p,q}$. Its distribution is the *iid* product of a Bernoulli-p distribution on finite dimensional sets, and is the *iid* product of a Bernoulli-q distribution on the tail-field events.[15] P satisfies the weak-law of large numbers over finite sequences with Bernoulli parameter p and satisfies the strong-law of large numbers

the same probability to each finite history of coin flips. Letting h_n denote a specific history of length n,

$$P^{5/10,\ 1/10}(h_n) = P^{5/10,\ 9/10}(h_n) = 2^{-n}.$$

But then $$P'(L^{1/10} \mid h_n) = P'(L^{9/10} \mid h_n) = \text{"} = P'(L^{1/10}) = P'(L^{9/10}),$$

for each possible history. That is, contrary to the strengthened convergence-to-the-truth result, in this modified P'-model, the agent is completely certain that her posterior probability for either of the two tail-field hypotheses, $L^{1/10}$ or $L^{9/10}$, is stationary at the prior value 1/2. Under the growing finite histories from each infinite sequence of coin flips, the posterior probability moves neither towards 0 nor toward 1. Within the P'-model, surely there is no convergence to the truth about these two tail-field events given increasing evidence from coin-flipping.[16]

Evidently, one aspect of what is unsettling about these finitely additive coin models is that the observed sequence of flips is entirely uninformative about the change point variable, *N*. No matter what the observed sequence, the agent's posterior distribution for *N* is her/his prior distribution for *N*, which is a purely finitely additive distribution assigning 0 probability to each possible integer value for *N*. It is not merely that this Bayesian agent cannot learn about the value of *N* from finite histories. Also, two such agents who have finitely additive coin models that disagree only on the tail-field parameter cannot use the shared evidence of the finite histories to induce a consensus about the tail-field N opinion. Peirce asserts that the scientific method for resolving such disputes wins over other rivals (e.g., *apriorism*, or the *method of tenacity*) by having the *Truth* (aka observable *Reality*) win out – by settling debates through an increasing sequence of observations from well designed experiments. With due irony, much of Peirce's proposal for letting *Reality* settle the intellectual dispute is embodied within personalist Bayesian methodology.[17] Here, we review some of those Bayesian resources regarding three aspects of *immodesty*.

One kind of epistemic *immodesty* is captured in a dogmatic credal state that is immune to

Not surprising then, as the community   increases its membership, the kind of consensus that is assured – the version of community-wide probabilistic merging that results from shared evidence – becomes weaker. So, one way to assess the epistemological "immodesty" of a credal state formulated with respect to a measurable space <$\boldsymbol{X}$, $\mathcal{B}$> is to identify the breadth of the community   of rival credal states that admits merging through increasing shared evidence from $\mathcal{B}$. For example, the agent who thinks each morning that it is highly probable that the world ends later that afternoon has an *immodest* attitude because there is only the isolated community of like-minded pessimists who can reconcile their views with commonplace evidence that is shared with the rest of us.

When the different opinions do not satisfy the requirement of mutual absolute continuity, the previous results do not apply directly. Instead, we modify an idea from Levi [1980, §13.5] so that different members of a community of investigators modify their individual credences (using convex combinations of rival credal states) in order *to give other views a hearing* and, in Peircean fashion, in order to allow increasing shared evidence to resolve those differences.

Let $I = \{i_1, \ldots\}$ serve as a finite or countably infinite index set, and let   = {P_i: i I} represent a community of investigators, each with her/his own countably additive credence function P_i on a common measurable space <$\boldsymbol{X}$, $\mathcal{B}$>. It may be that, pairwise, the elements of   are not even mutually absolutely continuous. In order to allow new evidence to resolve differences among the investigators' credences for elements of $\mathcal{B}$ (rather than trying, e.g., to preserve common judgments of conditional credal independence between pairs of elements of $\mathcal{B}$) each member of   shifts to a credal state by taking a mixture of each of the investigators' credal states: a "linear pooling" of those states. Specifically, for each i I, let 7! 8%{ . $_{ij}$: . $_{ij}$ > 0, assumptions for the Blackwell-Dubins' result (***) despite being self-centered. Depending upon the size of the community , using the replacement credal states {Q_i} results (1), (2), and (3) obtain.

We conclude this discussion of probabilistic merging with a reminder that merely finitely additive probability models open the door to reasoning to a foregone conclusion, Kadane Schervish, and Seidenfeld [1996], in a different sharp contrast with the P' model above to the almost sure asymptotic merging and convergence-to-the-truth results associated with countably additive probability models. Key to these asymptotic results for countably additive probabilities is the *Law of Iterated Expectations*.

Let X and Y be (bounded) random variables measurable with respect to a countably additive measure space <&, $\mathcal{B}$, P>. With [X] and [X | $Y = y$] denoting, respectively, the expectation of X and the conditional expectation of X, given the event $Y = y$, then

Law of Iterated Expectations [X] = [[X | Y]].

As Schervish et al. established [1984], each merely finitely (and not countably) additive probability defined on a

frequencies. In addition, the P-model fails Condition #1, which we understand is one of Belot's standards for *modesty*.

As we illustrate in Section 3, the conditional probabilities arising from sample paths of X. We denote elements of $_X$ as $y = <y_1, y_2, \ldots>$. $_X$ is a subset of 1

For each y X, define $\mathbf{5}_0(y) = 0$, and for $j > 0$, define

$$\mathbf{5}_j(y) = \begin{cases} \text{6>;<D< L M53+"N(0N) B F5E} & \text{if the minimum is finite,} \\ \text{7} & \\ \text{8 \%} & \text{if not.} \end{cases}$$

Let $B = \{y$ X : M5"N(< % for all j}, and let A = $X \setminus B$ = B^c 9 X.

Note that A is the set of sample paths each of which fails to visit at least one of the B_j sets in the order specified. Because we do not require that the sets B_j are nested, it is possible that the sequence reaches B_k for all $k > j$ without ever reaching B_j. Or the sequence could reach B_j before reaching B_{j-1} but not after.

Theorem A1: A is a meager set.

Proof: Write A = , $_j$ C_j, where $C_j = \{y$ X: M5"N(= %}. Then A is meager if and only if C_j is meager for every j. We prove that C_j is meager for every j by induction.

Start with $j = 1$. We have M5"N(= % *if and only if* y C_1 = $\bigcap$9)*+ {y X: y_n FP+}. To see that C_1 is meager, notice that QP+ = $\bigcup$9)*+ R), where

D_1 = X 0 . 24000 . 24 3 75 . 4 9 7 3 5650005000 / 3 00 . 2420 . 24462 0050 3 5421

ut Tf (y) Tj ET Q q 0 24 0 0R

References

Belot, G. (2013) *Bayesian Orgulity*. Phil. Sci. 80: 483-503.
Billingsley, P (1986) *Probability and Measure*, 2[nd]