The Topology of Statistical Verifiability

Konstantin Genin

Department of Philosophy
Carnegie Mellon University
Pittsburgh, Pennsylvania

konstantin.genin@gmail.com

Kevin T. Kelly

Department of Philosophy
Carnegie Mellon University
Pittsburgh, Pennsylvania

kk3n@andrew.cmu.edu

Topological models of empirical and formal inquiry are increasingly prevalent. They have emerged in such diverse fields as domain theory [1, 16], formal learning theory [18], epistemology and philosophy of science [10, 15, 8, 9, 2], statistics [6, 7] and modal logic [17, 4]. In those applications, open sets are typically interpreted as hypotheses deductively verifiable by true *propositional* information that rules out relevant possibilities. However, in statistical data analysis, one routinely receives random samples logically compatible with every statistical hypothesis. We bridge the gap between propositional and statistical data by solving for the *unique* topology on probability measures in which the open sets are exactly the *statistically verifiable* hypotheses. Furthermore, we extend that result to a topological characterization of learnability in the limit from statistical data.

1 Verifiability from Propositional Information

The results in this section appear in [5], [8], and [2], but we restate them here to clarify the intended analogy between propositional and statistical verifiability. Let W be a set of possible worlds, or possibilities one takes seriously, consistent with the background assumptions of inquiry. A proposition is identified with the set of worlds in which it is true, so propositions are subsets of W. Let P, Q be arbitrary propositions. Logical operations correspond to set-theoretic operations in the usual way: $P \cap Q$ is conjunction, $P \cup Q$ is disjunction, $P^c = W \setminus P$ is negation, and $P \subseteq Q$ is deductive entailment of Q by P. Finally, P is deductively valid iff $P = W$ and is deductively contradictory iff $P = \varnothing$.

In the propositional information setting, *information states* are propositions that rule out relevant possibilities. For every w in W, let $\mathscr{I}_w$ be the set of all information states true in w. It is assumed that $\mathscr{I}_w$ is non-empty (at worst, one receives the trivial information W). Furthermore, it is assumed that for each E, F in $\mathscr{I}_w$, there exists G in $\mathscr{I}_w$ such that $G \subseteq E \cap F$. The underlying idea is that a sufficiently diligent inquirer in w eventually receives information as strong as an arbitrary information state E in $\mathscr{I}_w$. Since that is true of both E and F, there must be true information as strong as $E \cap F$.

Example 1.1. *Let W be the set of all infinite binary sequences. Each world w determines an infinite sequence of observable outcomes. Let $w|_n$ be the initial segment of w of length n. Let $[w|_n]$ be the set of all worlds having $w|_n$ as an initial segment. Let $\mathscr{I}_w$ be the set of all $[w|_n]$ for every n. Think of the length of the initial segment observed as the "stage" of inquiry. There is exactly one such information state in w at every stage, and $[w|_n]$ is entailed by $[w|_m]$ for every $m \geq n$.*

Example 1.2. *Let W be the set of all real numbers. Think of the possible "stage-n" information states in w as the open intervals of width $1/2^n$ that contain w. Then $\mathscr{I}_w$ is the set of all intervals containing w of width $1/2^n$, for some natural number n. It follows that for every $E \in \mathscr{I}_w$ there is a stage n such that every stage-n information state in $\mathscr{I}_w$ entails E.*

Let $\mathscr{I} = \bigcup_w \mathscr{I}_w$, the set of all possible information states. It follows from the two assumptions on $\mathscr{I}_w$ that $\mathscr{I}$ is a topological basis. Therefore, the closure of topological basis $\mathscr{I}$ under union, denoted $\mathscr{T}$, is a topological space. We assume that $\mathscr{I}$ is countable, since any language in which the data are recorded is at most countably infinite. The elements of $\mathscr{T}$ are called *open sets*. The complements of open sets are called *closed sets*. A *clopen set* is both open and closed. A *locally closed set* is the intersection of an open and a closed set. Information state E *verifies* proposition H iff E entails H. The *interior* of a proposition H, denoted $\mathsf{int}(H)$, is the set of all worlds w, such that there is $E \in \mathscr{I}_w$ verifying H. Hence, $\mathsf{int}(H)$ is the set of worlds in which H is eventually verified by information. It is an elementary result that H is open iff $H = \mathsf{int}(H)$. The *closure* of H, denoted $\mathsf{cl}(H)$, is the set of all worlds in which H is compatible with all information, i.e. $\mathsf{int}(H^c)^c$. The boundary of H, denoted $\mathsf{bdry}(H)$, is defined as $\mathsf{cl}(H) \cap \mathsf{cl}(H^c)$. Every information state consistent with $\mathsf{bdry}(H)$ is consistent with both H and H^c.

A *method* is a function from information states to propositions. Method $L(\cdot)$ is *infallible* iff its output is always true, i.e. iff $w \in L(E)$ for all $E \in \mathscr{I}_w$.[1] Suppose that one desires to arrive at true belief concerning some proposition H without exposing oneself to the possibility of error. A *verifier* for H is an infallible method that converges to belief in H iff H is true. That is, $L(\cdot)$ is a verifier for H iff

V1. $L(\cdot)$ is infallible and

V2. $w \in H$ iff there is $E \in \mathscr{I}_w$ such that $L(F) \subseteq L(E) \subseteq H$ for all $F \in \mathscr{I}_w$ entailing E.

Say that H is *verifiable* iff there exists a verifier for H. Say that H is *refutable* iff its complement is verifiable, and that H is *decidable* iff H is both verifiable and refutable. For example, if you are observing a computation by an unknown program, it is verifiable that that the program will halt at some point, but it is not verifiable that it will never halt. In the setting of Example 1.1, it is verifiable that a zero will be observed at some stage, but not that it will be observed at every stage. Verifiability is fundamentally a topological concept.

Theorem 1.1. *Proposition H is verifiable iff H is open.*

Theorem 1.1 implies that if H is not open, then there is in general no error-avoiding method that arrives at true belief in H. Every method that converges to true belief in worlds in which H is never verified must leap beyond the information available, and expose itself to error thereby.

The infallibility requirement is too strict to allow for inductive learning that draws conclusions beyond the information provided. The following success criterion is less demanding. A *limiting verifier* for H is a method that converges to true belief in H iff H is true. That is, $L(\cdot)$ is a limiting verifier for H iff it satisfies V2. Say that H is *limiting verifiable* iff there exists a limiting verifier of H. In the setting of Example 1.1, no method verifies the constantly-zero hypothesis $\{000\ldots\}$, but that hypothesis is verified in the limit by the method that conjectures $\{000\ldots\}$ as long as it is not refuted by information. The following is a topological characterization of the propositions that are verifiable in the limit:

Theorem 1.2. *Proposition H is limiting verifiable iff H is a countable union of locally closed sets. If $\mathscr{T}$ is metrizable, then H is limiting verifiable iff H is a countable union of closed sets.*

Finally, an *empirical problem* is a countable partition $\mathscr{Q}$ of the worlds in W into a set of *answers*. For $w \in W$, write $\mathscr{Q}_w$ for the answer true in w. A method is a *solution* to $\mathscr{Q}$ iff it converges, on increasing information, to the true answer in $\mathscr{Q}$, i.e. iff for every $w \in W$, there exists $E \in \mathscr{I}_w$ such that $L(F) \subseteq \mathscr{Q}_w$ for all $F \in \mathscr{I}_w$ entailing E. A problem is *solvable* iff it has a solution.

---

[1]This is equivalent to requiring that the method's conclusions are deductively entailed by the data, i.e. that $E \subseteq L(E)$ for all information states E. For this reason, infallible methods are *deductive* and vice-versa.

Theorem 1.3. *Problem $\mathscr{Q}$ is solvable iff every answer is a countable union of locally closed sets.*

Theorems like 1.1, 1.2, and 1.3 constitute an exact correspondence between topology and learnability.

2   Verifiability from Statistical Information

There is a seeming gulf between propositional information and statistical samples. Propositional information literally rules out relevant possibilities. In sharp contrast, a random sample is often logically compatible with *every* possible probability distribution. We sidestep that fundamental difficulty by solving for the *unique* topology in which the open sets are precisely the *statistically verifiable* propositions, which provides an exact, statistical analogue of Theorem 1.1.

2.1   Samples and Worlds

A *sample space* $\mathfrak{S} = (\Omega, \mathscr{T})$ is a set of possible random samples Ω equipped with a topology $\mathscr{T}$ generated by a basis $\mathscr{I}$. The worlds in W assign probabilities to every set in $\mathscr{B}$, the Borel σ-algebra generated by the topology on $\mathfrak{S}$. The topology on the sample space reflects what is verifiable about the sample itself. As in the purely propositional setting, it is *verifiable* that sample ω lands in A iff A is open, and it is *decidable* whether sample ω falls into region A iff A is clopen. For example, suppose that region A is the closed interval $[1/2, \infty]$, and suppose that the sample ω happens to land right on the end-point $1/2$ of A. Suppose, furthermore, that given enough time and computational power, the sample ω can be specified to arbitrary, finite precision. But no finite degree of precision: $\omega \approx .50$; $\omega \approx .500$; $\omega \approx .5000$; ... suffices to determine that ω is truly in A. But the mere possibility of a sample hitting the boundary of A does not matter statistically, if the chance of obtaining such a sample is zero. A Borel set A for which $\mu(\mathrm{bdry}(A)) = 0$ is said to be *almost surely clopen (decidable) in μ.*[2] Borel set A is almost surely clopen iff it is almost surely clopen in every μ in W, and a collection of Borel sets $\mathscr{S}$ is almost surely clopen iff every element of $\mathscr{S}$ is almost surely clopen.

Example 2.1. *Consider the outcome of a single coin flip. The set Ω of possible outcomes is $\{H, T\}$. Since every outcome is decidable, the appropriate topology on the sample space is $\mathscr{T} = \{\varnothing, \{H\}, \{T\}, \{H, T\}\}$, the discrete topology on Ω. Let W be the set of all probability measures assigning a bias to the coin. Since every element of $\mathscr{T}$ is clopen, every element is also almost surely clopen.*

Example 2.2. *Consider the outcome of a continuous measurement. Then the sample space Ω is the set of real numbers. Let the basis $\mathscr{I}$ of the sample space topology be the usual interval basis on the reals. That captures the intuition that it is verifiable that the sample landed in some open interval, but it is not verifiable that it landed exactly on the boundary of an open interval. There are no nontrivial decidable (clopen) propositions in that topology. However, in typical statistical applications, W contains only probability measures μ that assign zero probability to the boundary of an arbitrary open interval. Therefore, every open interval E is almost surely decidable, i.e. $\mu(\mathrm{bdry}(E)) = 0$.*

Product spaces represent the outcomes of repeated sampling. Let I be an index set, possibly infinite. Let $(\Omega_i, \mathscr{T}_i)_{i \in I}$ be sample spaces, each with basis $\mathscr{I}_i$. Define the *product* $(\Omega, \mathscr{T})$ of the $(\Omega_i, \mathscr{T}_i)$ as follows: let Ω be the Cartesian product of the Ω_i; let $\mathscr{T}$ be the product topology, i.e. the topology in which the open sets are unions of Cartesian products $\times_i O_i$, where each O_i is an element of $\mathscr{T}_i$, and all but finitely

---

[2]A set that is almost surely clopen in μ is sometimes called a *continuity set* of μ.

many O_i are equal to Ω_i. When I is finite, the products of basis elements in $\mathscr{I}_i$ are the intended basis for $\mathscr{T}$. Let $\mathscr{B}$ be the σ-algebra generated by $\mathscr{T}$. Let μ_i be a probability measure on $\mathscr{B}_i$, the Borel σ-algebra generated by the $\mathscr{T}_i$. The *product measure* $\mu = \times_i \mu_i$ is the unique measure on $\mathscr{B}$ such that, for each $B \in \mathscr{B}$ expressible as a Cartesian product of $B_i \in \mathscr{B}_i$, where all but finitely many of the B_i are equal to Ω_i, $\mu(B) = \prod \mu_i(B_i)$. Let $\mu^{|I|}$ denote the $|I|$-fold product of μ with itself.

2.2 Statistical Tests

A statistical *method* is a measurable function from random samples to propositions over W.[3] A *test* of a statistical hypothesis $H \subseteq W$ is a statistical method $\psi : \Omega \to \{W, H^c\}$. Call $\psi^{-1}(W)$ the *acceptance region*, and $\psi^{-1}(H^c)$ the *rejection region* of the test.[4] The *power* of test $\psi(\cdot)$ is the worst-case probability that it rejects truly, i.e. $\inf_{\mu \in H^c} \mu[\psi^{-1}(H^c)]$. The *significance level* of a test is the worst-case probability that it rejects falsely, i.e. $\sup_{\mu \in H} \mu[\psi^{-1}(H^c)]$.

A test is *feasible in* μ iff its acceptance region is almost surely decidable in μ. Say that a test is *feasible* iff it is feasible in every world in W. More generally, say that a method is *feasible* iff the preimage of every element of its range is almost surely decidable in every world in W. Tests that are not feasible in μ are impossible to implement — as described above, if the acceptance region is not almost surely clopen in μ, then with non-zero probability, the sample lands on the boundary of the acceptance region, where one cannot decide whether to accept or reject. If one were to draw a conclusion at some finite stage, that conclusion might be reversed in light of further computation. Tests are supposed to *solve* inductive problems, not to generate new ones. Therefore we consider only feasible methods in the following development.

2.3 The Weak Topology

A sequence of measures $(\mu_n)_n$ *converges weakly* to μ, written $\mu_n \Rightarrow \mu$, iff $\mu_n(A) \to \mu(A)$ for every A almost surely clopen in μ. It is immediate that $\mu_n \Rightarrow \mu$ iff for every μ-feasible test $\psi(\cdot)$, $\mu_n(\psi \text{ rejects}) \to \mu(\psi \text{ rejects})$. It follows that no feasible test of $H = \{\mu\}$ achieves power strictly greater than its significance level. Furthermore, every feasible method that correctly infers H with high chance in μ, exposes itself to a high chance of error in "nearby" μ_n. It is a standard fact that one can topologize W in such a way that weak convergence is exactly convergence in the topology: the usual sub-basis is given by sets of the form $\{\nu : |\mu(A) - \nu(A)| < \varepsilon\}$, where A is almost surely clopen in μ.[5] That topology is called the *weak topology*. If $\mathfrak{S}$ is second countable and metrizable, then the weak topology on W is also second countable and metrizable, e.g. by the Prokhorov metric [3, Theorem 6.8]. When $\mathscr{I}$ is countable and almost surely clopen, the weak topology is generated in a particularly natural way.[6]

Lemma 2.1. *Suppose that $\mathscr{I}$ is a countable, almost surely clopen basis for W. Let $\mathscr{A}$ be the algebra generated by $\mathscr{I}$. Then the collection $\{\mu : \mu(A) \in (a, b)\}$ for $A \in \mathscr{A}$ and $a, b \in \mathbb{Q}$ is a countable sub-basis*

---

[3] The σ-algebra on the range of the method is assumed to be the power set.

[4] The acceptance region is $\psi^{-1}(W)$, rather than $\psi^{-1}(H)$, because failing to reject H licenses only the trivial inference W.

[5] Recall that a sequence μ_n converges to μ in a topology iff for every open set E containing μ, there is n_0 such that $\mu_n \in E$ for all $n \geq n_0$. If a topology is first countable, μ_n converge to μ in the topology iff μ is in the topological closure of the μ_n.

[6] That condition is satisfied, for example, in the standard case in which the worlds in W are Borel measures on $\mathbb{R}^n$, and all measures are absolutely continuous with respect to Lebesgue measure, i.e. when all measures have probability density functions, which includes normal, chi-square, exponential, Poisson, and beta distributions. It is also satisfied for discrete distributions like the binomial, for which the topology on the sample space is the discrete (power set) topology, so every acceptance zone is clopen and, hence, feasible. Naturally, it is satisfied in the particular cases of Examples 2.1 and 2.2.

for the weak topology.

That sub-basis for the weak topology has two fundamental advantages over the standard sub-basis. First, its closure under finite intersection is evidently a countable basis. Second, it is easy to show that the sub-basis elements are statistically verifiable. The following observations are easy consequences of the Lemma. In the setting of Example 2.1, the set of all $\{\mu : \mu(\{H\}) \in (a,b)\}$ for $a,b \in \mathbb{Q}$, assigning open intervals of biases for the coin, forms a sub-basis for the weak topology on W. In fact, it forms a basis. If μ is the world in which the bias of the coin is exactly .5 and μ_n is the world in which the bias is exactly $.5 + 1/2^n$, then the μ_n converge to μ in the weak topology.

3 Statistical Verifiability

In Section 1, proposition H was said to be verifiable iff there is an infallible method that converges on increasing information to H iff H is true. That condition implies that there is a method that achieves *every* bound on *chance* of error, and converges to H iff H is true.[7] In statistical settings, one cannot insist on such a high standard of infallibility. Instead, say that H is *verifiable in chance* iff for every bound on error, there is a method that achieves it, and that converges in probability to H iff H is true. The reversal of quantifiers expresses the fundamental difference between statistical and propositional verifiability and, hence, between statistical and propositional information. Say that a family $\{\lambda_n\}_{n\in\mathbb{N}}$ of feasible tests of H^c is an *α-verifier in chance* of $H \subseteq W$ iff for all $n \in \mathbb{N}$:

SV1. $\mu^n[\lambda_n^{-1}(H)] \leq \alpha$, for all $\mu \in H^c$ and

SV2. $\lim_{n\to\infty} \mu^n[\lambda_n^{-1}(H)] = 1$, for all $\mu \in H$.

Say that $H \subseteq W$ is *α-verifiable in chance* iff there is an α-verifier in chance of H. Say that $H \subseteq W$ is *verifiable in chance* iff H is α-verifiable in chance for every $\alpha > 0$.

The preceding definition only bounds the chance of error at each sample size. One might strengthen SV1 to the requirement that the overall chance of error be bounded, when H is false. Furthermore, one might also strengthen SV2 by requiring almost sure convergence to H, rather than mere convergence in probability in every measure in W. Say that a family $\{\lambda_n\}_{n\in\mathbb{N}}$ of feasible tests of $H^c \subseteq W$ is an *almost sure α-verifier* of H iff

SV3. $\sum_{n=1}^{\infty} \mu^n[\lambda_n^{-1}(H)] \leq \alpha$ for all $\mu \in H^c$ and

SV4. $\mu^\infty \left[\liminf_{n\to\infty} \lambda_n^{-1}(H) \right] = 1$ for all $\mu \in H$.

Say that $H \subseteq W$ is *almost surely α-verifiable* iff there is an almost sure α-verifer of H. Say that H is *almost surely verifiable* iff H is almost surely α-verifiable, for every $\alpha > 0$. Clearly, if H is almost surely verifiable, then H is verifiable in chance.

We now weaken the preceding two criteria of statistical verifiability to arrive at statistical notions of limiting verifiability. Say that a family $\{\lambda_n\}_{n\in\mathbb{N}}$ of feasible methods is a *limiting verifier in chance* of $H \subseteq W$ iff

---

[7]If for every $\varepsilon > 0$ your chance of error is less than ε, then your chance of error is zero: you are *almost surely* infallible.

1. $\mu \in H$ iff there is $H' \subseteq H$, s.t. $\lim\limits_{n \to \infty} \mu^n[\lambda_n^{-1}(H')] = 1$;

2. $\mu \notin H$ iff for all $H' \subseteq H$, $\lim\limits_{n \to \infty} \mu^n[\lambda_n^{-1}(H')] = 0$.

Say that $H \subseteq W$ is *limiting verifiable in chance* iff there is a limiting verifier in chance of H.

As before, there is an almost sure version of that success criterion. Say that a family $\{\lambda_n\}_{n \in \mathbb{N}}$ of feasible methods is a *limiting almost sure verifier* of $H \subseteq W$ iff

1. $\mu \in H$ iff there is $H' \subseteq H$, s.t. $\mu^\infty[\liminf\limits_{n \to \infty} \lambda_n^{-1}(H')] = 1$;

2. $\mu \notin H$ iff for all $H' \subseteq H$, $\mu^\infty[\limsup\limits_{n \to \infty} \lambda_n^{-1}(H')] = 0$.

Say that $H \subseteq W$ is *limiting a.s. verifiable* iff there is a limiting a.s. verifier of H.

Finally, there is a natural statistical analogue of solvability. Recall that an *empirical problem* is a countable partition $\mathcal{Q}$ of the worlds in W into a set of *answers*. Say that a family $\{\lambda_n\}_{n \in \mathbb{N}}$ of feasible methods is a *solution in chance* to $\mathcal{Q}$ iff for every $\mu \in W$, $\lim_{n \to \infty} \mu^n[\lambda_n^{-1}(\mathcal{Q}_\mu)] = 1$. Say that $\mathcal{Q}$ is *solvable in chance* iff there exists a solution in chance to $\mathcal{Q}$. A family $\{\lambda_n\}_{n \in \mathbb{N}}$ of feasible methods is an *almost sure solution* to $\mathcal{Q}$ iff for every $\mu \in W$, $\mu^\infty[\liminf_{n \to \infty} \lambda_n^{-1}(\mathcal{Q}_\mu)] = 1$. Furthermore, say that $\mathcal{Q}$ is *almost surely solvable* iff there exists an almost sure solution to $\mathcal{Q}$.

4 Results

Theorem 4.1 states that, for sample spaces with countable, almost surely clopen bases, verifiability in chance and almost sure verifiability are equivalent to being open in the weak topology. As promised in the introduction, that fundamental result lifts the topological perspective to inferential statistics.

Theorem 4.1. *Suppose that W is a set of Borel measures on $\mathfrak{S} = (\Omega, \mathcal{T})$, a metrizable sample space with countable, almost surely clopen basis $\mathcal{I}$. Then the following are equivalent:*

1. *$H \subseteq W$ is α-verifiable in chance for some $\alpha > 0$;*

2. *$H \subseteq W$ is almost surely verifiable;*

3. *$H \subseteq W$ is open in the weak topology.*

For an elementary application of the Theorem, consider, in the setting of Example 2.1, the sharp hypothesis that the bias of the coin is exactly .5. That hypothesis is almost surely refutable, but it is not almost surely verifiable. Since a topological space is determined uniquely by its open sets, Theorem 4.1 implies that the weak topology is the *unique* topology that characterizes statistical verifiability under the weak conditions stated in the antecedent of the theorem. Thus, under those conditions, the weak topology is not merely a convenient formal tool—it is *the* topology of statistical information.

Here is the promised statistical analogue of Theorem 1.2.

Theorem 4.2. *Suppose that W is a set of Borel measures on $\mathfrak{S} = (\Omega, \mathcal{T})$, a metrizable sample space with countable, almost surely clopen basis $\mathcal{I}$. Then the following are equivalent:*

1. *$H \subseteq W$ is limiting verifiable in chance;*

  2. *H ⊆ W is limiting almost surely verifiable;*

  3. *H ⊆ W is a countable union of closed sets in the weak topology.*

Finally, there is a natural statistical analogue of Theorem 1.3.

Theorem 4.3. *Suppose that W is a set of Borel measures on $\mathfrak{S} = (\Omega, \mathscr{T})$, a metrizable sample space with countable, almost surely clopen basis $\mathscr{I}$. Then the following are equivalent:*

  1. *$\mathscr{Q}$ is solvable in chance;*

  2. *$\mathscr{Q}$ is almost surely solvable;*

  3. *$\mathscr{Q}$ partitions W into countable unions of closed sets in the weak topology.*[8]

5   Related Work

Section 1 recapitulates foundational results in topological learning theory. Results stated in that section appear previously in [5], [8], and [2]. The theorems stated in section 4 are new, as far as we can tell. In statistical terminology, our Theorem 4.1 provides necessary and sufficient conditions for the existence of a Chernoff consistent test. Although there is extensive statistical work on pointwise consistent hypothesis testing, we are unaware of any topological result analogous to Theorem 4.1. The closest work is [14], where a topological characterization is given for consistent hypothesis testing of ergodic processes with samples from a discrete, finite alphabet. That result is incomparable with our own, because, although our work is done in the i.i.d setting, we allow samples to take values in an arbitrary, separable metric space. Furthermore, the topology employed in [14] is not the weak topology, but the topology of distributional distance. The existence of uniformly consistent tests is investigated topologically in [7], where some sufficient conditions are given. Limiting statistical solvability, or *discernability*, as it is known in the statistical literature, has been investigated topologically in [6] and [11]. The results of [6] are generalized to ergodic processes in [12]. Although the setting is slightly different, our Theorem 4.3 gives a simpler back-and-forth condition than the one given in [6] and is arrived at more systematically, by building on the fundamental Theorem 4.1. The weak topology is used in [6], but our Theorem 4.1 shows that the weak topology is the *unique* topology for which the open sets are exactly the statistically verifiable propositions. Our result shows, therefore, that the weak topology is more than just a convenient technical device.

6   Conclusion

This note lifts the topological perspective on empirical inquiry to statistics. In the deductive setting, open sets are deductively verifiable by true, propositional information. Theorem 4.1 exhibits a topology on probability measures in which the open sets are exactly the propositions statistically verifiable from random samples. In the deductive setting, learnability in the limit receives an elegant topological characterization [2, 8]. Theorems 4.2 and 4.3 provide analogous topological characterizations of learnability in the limit from statistical data. In light of those fundamental bridge results, we expect many of the streamlined insights of formal learning theory to apply literally to the concrete statistical problems that

---

[8]A similar result is proven in [6, Theorem 2] under different conditions. Dembo and Peres do not require their methods to be feasible, so Theorem 4.2 does not straightforwardly generalize their result. It is not difficult to reprove Theorem 4.2 without that requirement to obtain a generalization of the result in [6].

arise in statistics and machine learning. Of particular interest is the learning theoretic vindication of Ockham's razor, developed topologically in [8], and [9].

References

[1] Samson Abramsky & Achim Jung (1994): *Domain theory*. Handbook of logic in computer science 3, pp. 1–168.

[2] Alexandru Baltag, Nina Gierasimczuk & Sonja Smets (2015): *On the Solvability of Inductive Problems: a study in epistemic topology*. In: Proceedings of the Fifteenth Conference on Theoretical Aspects of Rationality and Knowledge, doi:10.4204/EPTCS.215.7.

[3] Patrick Billingsley (1999): *Convergence of Probability Measures*. John Wiley & Sons, doi:10.1002/9780470316962.

[4] Adam Bjorndahl (2013): *Subset space public announcement logic revisited*. arXiv preprint arXiv:1302.4009.

[5] Matthew de Brecht & Akihiro Yamamoto (2009): *Interpreting Learners as Realizers for Σ_2^0-Measurable Functions*. (Manuscript).

[6] Amir Dembo & Yuval Peres (1994): *A topological criterion for hypothesis testing*. The Annals of Statistics, pp. 106–117, doi:10.1214/aos/1176325360.

[7] Mikhail Ermakov (2013): *On distinguishability of hypotheses*. arXiv preprint arXiv:1308.4295.

[8] Konstantin Genin & Kevin T. Kelly (2015): *Theory Choice, Theory Change, and Inductive Truth-Conduciveness*. In: Proceedings of the fifteenth conference on Theoretical Aspects of Rationality and Knowledge.

[9] Konstantin Genin & Kevin T. Kelly (forthcoming): *Learning, Theory Choice, and Belief Revision*. Studia Logica.

[10] Kevin T. Kelly (1996): *The Logic of Reliable Inquiry*. Oxford University Press.

[11] Sanjeev R Kulkarni & Ofer Zeitouni (1995): *A general classification rule for probability measures*. The Annals of Statistics, pp. 1393–1407, doi:10.1214/aos/1176324714.

[12] Andrew B. Nobel (2006): *Hypothesis testing for families of ergodic processes*. Bernoulli, pp. 251–269, doi:10.3150/bj/1145993974.

[13] Kalyanapuram Rangachari Parthasarathy (1967): *Probability measures on metric spaces*. 352, American Mathematical Soc.

[14] Daniil Ryabko (2011): *Learnability in Problems of Sequential Inference*. Ph.D. thesis, Université des Sciences et Technologie de Lille-Lille I.

[15] Oliver Schulte & Cory Juhl (1996): *Topology as epistemology*. The Monist 79(1), pp. 141–147, doi:10.5840/monist19967916.

[16] Steven Vickers (1996): *Topology Via Logic*. Cambridge University Press.

[17] Yì N Wáng & Thomas Ågotnes (2013): *Subset space public announcement logic*. In: Logic and Its Applications, Springer, pp. 245–257, doi:10.1007/978-3-642-36039-8_22.

[18] Akihiro Yamamoto & Matthew de Brecht (2010): *Topological Properties of Concept Spaces (full version)*. Information and Computation 208(4), pp. 327–340, doi:10.1016/j.ic.2009.08.001.

A   Proofs and Lemmas

A.1   Deductive Verifiability

Proof of Theorem 1.1. Right to left. Suppose that H is open, and that H is true in w. Let $L(E) = H$ if E entails H, and let $L(E) = W$ otherwise. Since H is a union of information states, there is an information state F true in w that entails H. Therefore, $L(F) = H$. Furthermore, for any information state G true in w, we have that $L(G \cap F) = H$. So L converges to true belief in H. Furthermore, if $E \in \mathscr{I}_w$ then either $w \in E \subseteq H = L(E)$, or $w \in W = L(E)$, so L avoids error in all worlds. *Left to right.* Suppose that H is not open. Then H is true in some w, such that for all information E true in w, E does not entail H, i.e. there is $w \in H \cap \mathrm{bdry}(H)$. Suppose, for contradiction, that L verifies H. Then $L(F) \subseteq H$, for some F true in w. But, by assumption, there is $v \in F \cap H^{\mathsf{c}}$. So L does not avoid error in v. $\qquad\square$

Proof of Theorem 1.2. Left to right. Suppose that $L(\cdot)$ is a limiting verifier of H. Let

$$\mathscr{T} = \{E \in \mathscr{I} : L(E) \subseteq H\}.$$

For each $E \in \mathscr{T}$, let $\mathscr{D}_E = \{F \in \mathscr{I} : F \subseteq E \text{ and } L(F) \nsubseteq L(E)\}$, and let $E' = \bigcup \mathscr{D}_E$. We claim that:

$$H = \bigcup_{E \in \mathscr{T}} E \setminus E'.$$

To prove the claim, $w \in H$ iff there is $E \in \mathscr{I}_w$ such that for all information states $F \subseteq E$, $L(F) \subseteq L(E) \subseteq H$ iff there is $E \in \mathscr{T}$ such that $w \in E \setminus E'$. Since $\mathscr{T} \subseteq \mathscr{I}$, and $\mathscr{I}$ is countable, H is expressed as a countable union of locally closed sets. If the topology is metrizable, every open set — and therefore every locally closed set — can be expressed as a countable union of closed sets. *Right to left.* Every countable union of locally closed sets can be expressed as a disjoint union of locally closed sets [2, Proposition 3]. Let $H = \sqcup_{i=1}^{\infty} O_i \setminus O_i'$ be a disjoint union, for O_i, O_i' open. Let $\sigma(E)$ be the least i such that $E \subseteq O_i$ and $E \nsubseteq O_i'$, if such an i exists, and let $\sigma(E) = \omega$ otherwise. Let $L(E) = O_{\sigma(E)} \setminus O_{\sigma(E)}'$ if $\sigma(E) < \omega$, and let $L(E) = W$ otherwise. Suppose that $w \in H$. Let k be the least integer such that $w \in O_k \setminus O_k'$. Then for $j < k$, either $w \notin O_j$ or $w \in O_j'$. For each $j < k$, let E_j be an information state true in w such that $E_j \subseteq O_j'$, if $w \in O_j'$, and let $E_j = W$, otherwise. Let E_k be an information state true in w that entails O_k. Finally, let $E = \bigcap_{j \leq k} E_k$. Then $L(F) = O_k \setminus O_k'$, for every $F \in \mathscr{I}_w$ such that $F \subseteq E$. Finally, suppose that $w \notin H$. Then for each j, either $w \notin O_j$ or $w \in O_j'$. Suppose that $L(E) = O_i \setminus O_i'$, for some i. Then $w \in O_i'$. Let F be an information state true in w and entailing O_i'. Then $L(E \cap F) \nsubseteq L(E)$, because $L(E \cap F)$ is either $W \nsubseteq O_i \setminus O_i'$, or $L(E \cap F)$ is some $O_j \setminus O_j'$, which was chosen to be disjoint from $O_i \setminus O_i'$. $\qquad\square$

Proof of Theorem 1.3. See the proof of Theorem 2 in [8]. $\qquad\square$

A.2   The Statistical Setting

A.2.1   The Sample Space

The following Lemma states that is always feasible to perform logical operations (e.g. $\wedge$, $\vee$, and $\neg$) on feasible tests.

Lemma A.1 (Lemma 6.4 [13])**.** *The almost surely clopen sets in μ, denoted $\mathscr{C}(\mu)$, form an algebra.*

Proof of Lemma A.1. One has that $\Omega \in \mathscr{C}(\mu)$, since $\mathrm{bdry}(\Omega) = \varnothing$. Moreover, $\mathscr{C}(\mu)$ is closed under complement, since $\mathrm{bdry}(A) = \mathrm{bdry}(A^{\mathsf{c}})$. Furthermore, since $\mathrm{bdry}(A \cup B) \subseteq \mathrm{bdry}(A) \cup \mathrm{bdry}(B)$, it follows that if $A, B \in \mathscr{C}(\mu)$, then $\mu(\mathrm{bdry}(A \cup B)) \leq \mu(\mathrm{bdry}(A) \cup \mathrm{bdry}(B)) \leq \mu(\mathrm{bdry}(A)) + \mu(\mathrm{bdry}(B)) = 0$. Therefore, $\mathscr{C}(\mu)$ is closed under finite union as well. $\qquad\square$

Hypothesis tests are often constructed to reject if the number of samples landing in a particular region exceeds some threshold. The following lemma states that such a test is μ-feasible, if the region is almost surely clopen in μ.

Lemma A.2. *Suppose that A is almost surely clopen in μ. Then:*

$$\left\{ (\omega_1, \ldots, \omega_n) : \sum_{i=1}^{n} \mathbb{1}[\omega_i \in A] \geq k \right\}$$

is almost surely clopen in μ^n, for $n \geq 1$, and $k \in \{0, \ldots n\}$.

Proof of Lemma A.2. Let $L_1, L_2, \ldots, L_{n\mathsf{C}k}$ enumerate all k-element subsets of $\{1, 2, \ldots, n\}$. Then

$$\{(\omega_1, \ldots, \omega_n) : \sum_{i=1}^{n} \mathbb{1}[\omega_i \in A] \geq k\} = \bigcup_{i=1}^{n\mathsf{C}k} \times_{j=1}^{n} B_{ij},$$

where $B_{ij} = A$ if $j \in L_i$, and $B_{ij} = \Omega$ otherwise. Since the almost surely clopen sets in μ^n form an algebra, it suffices to show that $\times_{j=1}^{n} B_{ij}$ is an almost surely clopen set in μ^n. Argue by induction on n. If $n = 1$, then $\times_{j=1}^{n} B_{ij}$ is either A or Ω, which are both almost surely clopen sets in μ. For the inductive step, note that $\mathsf{bdry}(\times_{j=1}^{n+1} B_{ij}) \subseteq \mathsf{bdry}(\times_{j=1}^{n} B_{ij}) \cup \mathsf{bdry}(B_{i,n+1})$. By the induction hypothesis, $\mu^{n+1}(\mathsf{bdry}(\times_{j=1}^{n+1} B_{ij})) \leq \mu^n(\mathsf{bdry}(\times_{j=1}^{n} B_{ij})) + \mu(\mathsf{bdry}(B_{i,n+1})) = 0$. $\square$

A.2.2 The Weak Topology

Billingsley [3] proves the following result about the product space:

Lemma A.3 (Theorem 2.8). *If $\mathscr{T} = \mathscr{T}' \times \mathscr{T}''$ is metrizable and second-countable, then $\mu_n' \times \mu_n'' \Rightarrow \mu' \times \mu''$ iff $\mu_n' \Rightarrow \mu'$ and $\mu_n'' \Rightarrow \mu''$.*

Lemma A.3 entails that the product map $\mu \mapsto \mu^n$ is sequentially continuous, and therefore, continuous, for all natural n.[9] Billingsley [3] also proves the following useful lemma:

Lemma A.4 (Theorem 2.2). *Suppose that $\mathscr{A} \subseteq \mathscr{B}$ is a π-system[10] and that every open set is a countable union of $\mathscr{A}$ sets. If $\mu_n(A) \to \mu(A)$ for every A in $\mathscr{A}$, then $\mu_n \Rightarrow \mu$.*

The following is a consequence of Lemma A.4.

Lemma A.5. *Suppose that $\mathscr{I}$ is a countable, almost surely clopen basis for W. Then the collection $\{\mu : \mu(A) > b\}$ for $A \in \mathscr{A}$ and $b \in \mathbb{Q}$ is a countable sub-basis for the weak topology.*

Proof of Lemma A.5. It is sufficient to show that $\mu_n \Rightarrow \mu$ iff the μ_n converge to μ in the topology generated by the sub-basis. *Left to right.* Suppose $\mu_n \Rightarrow \mu$. Let E be open in the topology generated by the sub-basis. Suppose μ lies in E. Then there is a basic open set:

$$B = \bigcap_{i=1}^{k} \{\mu : \mu(A_i) > b_i\},$$

---

[9]Continuity is relative to the weak topologies on W and $W^n(\mathfrak{S}) = \{\mu^n : \mu \in W\}$. Recall that a function f is sequentially continuous if whenever a sequence (x_n) converges to a limit x, the sequence $f(x_n)$ converges to $f(x)$. In first-countable spaces, sequential continuity is equivalent to continuity.

[10]$\mathscr{A}$ is a π-system iff $A \cap B \in \mathscr{A}$ whenever $A, B \in \mathscr{A}$.

such that $\mu \in B \subseteq E$. Since $\mathscr{I}$ is feasible for W, $\mu_n(B_i) \to \mu(B_i)$ for each i. Therefore, there exists n_i such that $\mu_n \in \{\mu : \mu(A_i) > b_i\}$ for all $n \geq n_i$. Letting $m = \max\{n_1, \ldots, n_k\}$, it follows that $\mu_n \in B \subseteq E$ for all $n \geq m$. Therefore, the μ_n converge to μ in the topology generated by the sub-basis. *Right to left.* Suppose that the μ_n converge to μ in the topology generated by the sub-basis. Note that since $\mathscr{A}$ is an algebra, the collection $\{\mu : \mu(A) \in (a,b)\}$ for $A \in \mathscr{A}$, and $a,b \in \mathbb{Q}$ generates the same topology.[11] Let $A_1, A_2, \ldots$ enumerate the elements of $\mathscr{A}(\mathscr{I})$. Let $l_{ij} < \mu(A_i) < r_{ij}$ be rationals lying in $(\mu(A_i) - 1/j, \mu(A_i) + 1/j)$. Let Z_{ij} denote the sub-basis element $\{\nu : \nu(A_i) \in (l_{ij}, r_{ij})\}$. Let f be a surjective function from $\mathbb{N}$ to $\mathbb{N} \times \mathbb{N}$. Let $U_k = Z_{f(k)}$. By assumption, for every $m \geq 1$, there is n_0 such that the μ_n lie in $\cap_{k=1}^m U_k$, for all $n \geq n_0$. So $\mu_n(A) \to \mu(A)$, for every $A \in \mathscr{A}$. By Lemma A.4, $\mu_n \Rightarrow \mu$. $\square$

Proof of Lemma 2.1. Immediate corollary of Lemma A.5. $\square$

A.3   Statistical Verifiability

First, a useful lemma.

Lemma A.6. *The almost surely verifiable propositions are closed under finite conjunctions, and countable disjunctions.*

Proof of Lemma A.6. Suppose that A_1, A_2 are a.s. verifiable. Let $\alpha > 0$. Let $\{\lambda_n^i\}_{n \in \mathbb{N}}$ be such that $\{\lambda_n^i\}_{n \in \mathbb{N}}$ is an a.s. α-verifier for A_i. Let $\lambda_n(\vec{\omega}) = A_1 \cap A_2$ if $\lambda_n^i(\vec{\omega}) = A_i$, for $i \in \{1,2\}$. By Lemma A.1, λ_n is feasible, for each $\mu \in W$, $n \in \mathbb{N}$. Suppose that $\mu \in A_1 \cap A_2$. Then:

$$\mu^\infty \left[\liminf_{n \to \infty} \lambda_n^{-1}(A_1 \cap A_2) \right] =$$
$$= \mu^\infty \left[\liminf_{n \to \infty} (\lambda_n^1)^{-1}(A_1) \cap (\lambda_n^2)^{-1}(A_2) \right]$$
$$= 1 - \mu^\infty \left[\limsup_{n \to \infty} (\lambda_n^1)^{-1}(W) \cup (\lambda_n^2)^{-1}(W) \right]$$
$$= 1 - \mu^\infty \left[\limsup_{n \to \infty} (\lambda_n^1)^{-1}(W) \cup \limsup_{n \to \infty} (\lambda_n^2)^{-1}(W) \right]$$
$$\geq 1 - \mu^\infty \left[\limsup_{n \to \infty} (\lambda_n^1)^{-1}(W) \right] - \mu^\infty \left[\limsup_{n \to \infty} (\lambda_n^2)^{-1}(W) \right]$$
$$= -1 + \mu^\infty \left[\liminf_{n \to \infty} (\lambda_n^1)^{-1}(A_1) \right] + \mu^\infty \left[\liminf_{n \to \infty} (\lambda_n^2)^{-1}(A_2) \right]$$
$$= 1.$$

Suppose that $\mu \notin A_1 \cap A_2$. Without loss of generality, suppose $\mu \notin A_1$. Then:

$$\sum_{n=1}^\infty \mu^\infty \left[\lambda_n^{-1}(A_1 \cap A_2) \right] =$$
$$= \sum_{n=1}^\infty \mu^\infty \left[(\lambda_n^1)^{-1}(A_1) \cap (\lambda_n^2)^{-1}(A_2) \right]$$
$$\leq \sum_{n=1}^\infty \mu^\infty \left[(\lambda_n^1)^{-1}(A_1) \right] \leq \alpha.$$

---

[11] Notice that $\{\mu : \mu(A) \in (a,b)\} = \{\mu : \mu(A) > a\} \cap \{\mu : \mu(A^c) > 1 - b\}$.

To show that the a.s. verifiable propositions are closed under countable union, suppose that $A_1, A_2, \ldots$ are a.s. verifiable. For $i \in \mathbb{N}$, let $\{\lambda_n^i\}_{n \in \mathbb{N}}$ be an a.s. α_i-verifier for A_i with $\alpha_i = \alpha/2^i$. Let $\lambda_n(\vec{\omega}) = \bigcup_{i=1}^{\infty} A_i$ if $\lambda_n^i(\vec{\omega}) = A_i$ for some $i \in \{1, \ldots, n\}$, and let $\lambda_n(\vec{\omega}) = W$ otherwise. By Lemma A.1, λ_n is feasible for each $\mu \in W, n \in \mathbb{N}$. Suppose that $\mu \in \bigcup_{i=1}^{\infty} A_i$. Then there exists $j \in \mathbb{N}$ such that $\mu \in A_j$. Furthermore:

$$\mu^{\infty}\left[\liminf_{n \to \infty} \lambda_n^{-1}(\cup_{i=1}^{\infty} A_i)\right] = \mu^{\infty}\left[\liminf_{n \to \infty} \cup_{k \leq n}(\lambda_n^k)^{-1}(A_k)\right]$$

$$\geq \mu^{\infty}\left[\liminf_{n \to \infty}(\lambda_n^j)^{-1}(A_j)\right] = 1.$$

Suppose that $\mu \notin \cup_{i=1}^{\infty} A_i$. Then:

$$\sum_{n=1}^{\infty} \mu^{\infty}\left[\lambda_n^{-1}(\cup_{i=1}^{\infty} A_i)\right] = \sum_{n=1}^{\infty} \mu^{\infty}\left[\cup_{k=1}^{n}(\lambda_n^k)^{-1}(A_k)\right]$$

$$\leq \sum_{n=1}^{\infty} \mu^{\infty}\left[\cup_{k=1}^{\infty}(\lambda_n^k)^{-1}(A_k)\right]$$

$$\leq \sum_{k=1}^{\infty} \sum_{n=1}^{\infty} \mu^{\infty}\left[(\lambda_n^k)^{-1}(A_k)\right]$$

$$\leq \sum_{k=1}^{\infty} \alpha/2^k = \alpha.$$

$\square$

Proof of Theorem 4.1. 1 implies 3. Suppose, for contradiction, that H is not open, but that $\{\lambda_n\}_{n \in \mathbb{N}}$ is an α-verifier in chance for H. Let $\mu \in H \cap \mathrm{bdry}H$. Then there is a sequence of μ_n in H^c such that $\mu_n \Rightarrow \mu$. Since $\{\lambda_n\}_{n \in \mathbb{N}}$ is a verifier for H, there is a sample size k such that $\mu^k(\lambda_k^{-1}(H)) > \alpha + \varepsilon$. By Lemma A.3, $\mu_n^k(\lambda_k^{-1}(H)) \to \mu^k(\lambda_k^{-1}(H))$. So there is a $\mu_m \in H^c$ such that $\mu_m^k(\lambda_k^{-1}(H)) > \alpha$. Contradiction.

3 implies 2. By Lemmas 2.1, A.6, it is sufficient to show that every element of the sub-basis:

$$\{\{\mu : \mu(A) > b\} : A \in \mathscr{A}(\mathscr{I}), b \in \mathbb{Q}\}$$

is a.s. verifiable. Let $B \in \mathscr{A}(\mathscr{I})$, and let $H = \{\mu : \mu(B) > b\}$, for some $b \in [0,1] \cap \mathbb{Q}$. Define the indicator random variable $\mathbb{1}_B : \Omega \to \{0,1\}$ by $\mathbb{1}_B(\omega) = 1$ if $\omega \in B$, otherwise $\mathbb{1}_B = 0$. Letting $t_n = \sqrt{\frac{1}{2n} \ln(\pi^2 n^2/6\alpha)}$, it follows from Hoeffding's inequality that:

$$\mu^n\left[\sum_{i=1}^{n} \mathbb{1}_B(\omega_i) \geq n(\mu(B) + t_n)\right] \leq \frac{6\alpha}{\pi^2 n^2}.$$

Let $\lambda_n(\vec{\omega}) = H$ if $\sum_{i=1}^{n} \mathbb{1}_B(\omega_i) \geq \lceil n(b + t_n) \rceil$, and let $\lambda_n(\vec{\omega}) = W$ otherwise. By Lemmas A.2, λ^n is

feasible for all μ^n. If $\mu \notin H$, then $b \geq \mu(B)$ and:

$$
\begin{aligned}
\sum_{n=1}^{\infty} \mu^n \left[\lambda_n^{-1}(H)\right] &= \sum_{n=1}^{\infty} \mu^n \left[\sum_{i=1}^{n} \mathbb{1}_B(\omega_i) \geq \lceil n(b+t_n)\rceil\right] \\
&\leq \sum_{n=1}^{\infty} \mu^n \left[\sum_{i=1}^{n} \mathbb{1}_B(\omega_i) \geq n(\mu(B)+t_n)\right] \\
&\leq \sum_{n=1}^{\infty} \frac{6\alpha}{\pi^2 n^2} = \alpha.
\end{aligned}
$$

Furthermore, if $\mu \in H$, then since $\frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_B(\omega_i) \overset{a.s.}{\to} \mathbb{E}[\mathbb{1}_B] = \mu(B)$, by the strong law of large numbers, and $t_n \to 0$, we have that $\frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_B(\omega_i) - t_n \overset{a.s.}{\to} \mu(B)$. Therefore, $\mu^{\infty}\left[\liminf_{n\to\infty} \lambda_n^{-1}(H)\right] = 1$, as required.

2 implies 1. Immediate from the definitions. □

Proof of Theorem 4.2. 1 entails 3. Suppose that $\{\lambda_n\}_{n\in\mathbb{N}}$ is a limiting verifier in chance of H. For every $H' \in \mathrm{rng}(\lambda_n)$, let:

$$
\begin{aligned}
\mathrm{trig}_n(H') &= \{\mu : \mu^n[\lambda_n^{-1}(H')] > \alpha\}; \\
\mathrm{def}_n(H') &= \cup_{m>n}\{\mu : \mu^n[\lambda_n^{-1}(H')^c] > 1-\alpha\}.
\end{aligned}
$$

Lemma A.3 and the feasibility of the λ_n entail that $\mathrm{trig}_n(H')$ and $\mathrm{def}_n(H')$ are both open in the weak topology. Let $\alpha \in (0,1)$. We claim that:

$$
H = \bigcup_{n=1}^{\infty} \bigcup_{H' \in \mathrm{rng}(\lambda_n)} \mathrm{trig}_n(H') \setminus \mathrm{def}_n(H').
$$

Observe that $\nu \in H$ iff there is $H' \subseteq H$ and $m \in \mathbb{N}$, such that for all $n \geq m$, $\nu^n[\lambda_n^{-1}(H')] > \alpha$ iff $\nu \in \mathrm{trig}_n(H') \setminus \mathrm{def}_n(H')$. Therefore, proposition H is a countable union of locally closed sets. Since the weak topology is metrizable, every open set — and therefore every locally closed set — can be expressed as a countable union of closed sets.

3 entails 2. Suppose that $H = \cup_{i=1}^{\infty} C_i$ is a countable union of closed sets. By theorem 4.1, for each C_i, there exists an a.s. statistical α-verifier of its complement $\{\psi_n^i\}_{n\in\mathbb{N}}$. Let $\lambda_n(\vec{\omega}) = C_j$, where j is the least integer in $1,\ldots,n$ such that $\psi_n^j(\vec{\omega}) = W$, if such a j exists. Otherwise, let $\lambda_n(\vec{\omega}) = W$. Suppose that

$\mu \in H$, and that k is the least integer such that $\mu \in C_k \subseteq H$. Then:

$$\mu^{\infty}\left[\liminf_{n\to\infty}\lambda_n^{-1}(C_k)\right] =$$

$$= \mu^{\infty}\left[\liminf_{n\to\infty}\bigcap_{j<k}(\psi_n^j)^{-1}(C_j^{\mathsf{c}})\cap(\psi_n^k)^{-1}(W)\right]$$

$$= 1 - \mu^{\infty}\left[\limsup_{n\to\infty}\bigcup_{j<k}(\psi_n^j)^{-1}(W)\cup(\psi_n^k)^{-1}(C_k^{\mathsf{c}})\right]$$

$$= 1 - \mu^{\infty}\left[\bigcup_{j<k}\limsup_{n\to\infty}(\psi_n^j)^{-1}(W)\cup\limsup_{n\to\infty}(\psi_n^k)^{-1}(C_k^{\mathsf{c}})\right]$$

$$\geq 1 - \sum_{j<k}\mu^{\infty}\left[\limsup_{n\to\infty}(\psi_n^j)^{-1}(W)\right] - \mu^{\infty}\left[\limsup_{n\to\infty}(\psi_n^k)^{-1}(C_k^{\mathsf{c}})\right].$$

Since each $\{\psi_n^j\}_{n\in\mathbb{N}}$ is an a.s. α-verifier of C_j, and $\mu \notin \cup_{j<k}C_j$, it follows that $\mu^{\infty}\left[\limsup_{n\to\infty}(\psi_n^j)^{-1}(W)\right] = 0$ for $j < k$. Therefore,

$$\mu^{\infty}\left[\liminf_{n\to\infty}\lambda_n^{-1}(C_k)\right] \geq 1 - \mu^{\infty}\left[\limsup_{n\to\infty}(\psi_n^k)^{-1}(C_k^{\mathsf{c}})\right].$$

Finally, since $\sum_{n=1}^{\infty}\mu^{\infty}\left[(\psi_n^k)^{-1}(C_k^{\mathsf{c}})\right] \leq \alpha$, by the Borel-Cantelli lemma,

$$\mu^{\infty}\left[\limsup_{n\to\infty}(\psi_n^k)^{-1}(C_k^{\mathsf{c}})\right] = 0,$$

and $\mu^{\infty}\left[\liminf_{n\to\infty}\lambda_n^{-1}(C_k)\right] = 1$, as required. Suppose that $\mu \notin H$. Then, for each C_k,

$$\mu^{\infty}[\limsup_{n\to\infty}\lambda_n^{-1}(C_k)] \leq \mu^{\infty}[\limsup_{n\to\infty}(\psi_n^k)^{-1}(W)] = 0.$$

2 entails 1. Immediate from the definitions. $\qquad\square$

Proof of Theorem 4.3. 1 entails 3. Suppose that the indexed set $\{\lambda_n\}_{n\in\mathbb{N}}$ is a solution in chance to $\mathcal{Q}$, and that A is an answer to $\mathcal{Q}$. Then, $\{\lambda_n\}_{n\in\mathbb{N}}$ is a limiting verifier in chance of A. Therefore, by Theorem 4.2, A is a countable union of closed sets in the weak topology.

3 entails 2. Let $A_1, A_2, \ldots$ enumerate the answers to $\mathcal{Q}$. Suppose that each answer is a countable union of closed sets, i.e. that each A_j in $\mathcal{Q}$ can be expressed as a countable union of closed sets $\cup_{i=1}^{\infty}C_{ij}$. By Theorem 4.1, for each C_{ij}, there exists an almost sure α-verifier $\{\psi_n^{ij}\}_{n\in\mathbb{N}}$ of its complement. Let f be a surjective function from $\mathbb{N}$ to $\mathbb{N}\times\mathbb{N}$. Let $\lambda_n(\vec{\omega}) = \mathcal{Q}(C_{f(k)})$, where k is the least natural number in $\{1,\ldots,n\}$ such that $\psi_n^{f(k)}(\vec{\omega}) = W$, if such a k exists, and let $\lambda_n(\vec{\omega}) = W$, otherwise. It is easy to see that each λ_n is feasible by Lemma A.1. Suppose that $\mu \in W$. Claim: $\mu^{\infty}[\liminf_{n\to\infty}\lambda_n^{-1}(\mathcal{Q}_\mu)] = 1$. To

establish the claim, let k be the least natural number such that $\mu \in C_{f(k)}$. Then:

$$\mu^\infty \left[\liminf_{n\to\infty} \lambda_n^{-1}(\mathscr{Q}_\mu) \right] \geq$$

$$\geq \mu^\infty \left[\liminf_{n\to\infty} \bigcap_{i<k} \left(\psi_n^{f(i)} \right)^{-1} \left(C_{f(i)}^{\mathsf{c}} \right) \cap \left(\psi_n^{f(k)} \right)^{-1} (W) \right]$$

$$= 1 - \mu^\infty \left[\limsup_{n\to\infty} \bigcup_{i<k} \left(\psi_n^{f(i)} \right)^{-1} (W) \cup \left(\psi_n^{f(k)} \right)^{-1} \left(C_{f(k)}^{\mathsf{c}} \right) \right]$$

$$= 1 - \mu^\infty \left[\bigcup_{i<k} \limsup_{n\to\infty} \left(\psi_n^{f(i)} \right)^{-1} (W) \cup \limsup_{n\to\infty} \left(\psi_n^{f(k)} \right)^{-1} \left(C_{f(k)}^{\mathsf{c}} \right) \right]$$

$$\geq 1 - \sum_{i<k} \mu^\infty \left[\limsup_{n\to\infty} \left(\psi_n^{f(i)} \right)^{-1} (W) \right] - \mu^\infty \left[\limsup_{n\to\infty} \left(\psi_n^{f(k)} \right)^{-1} \left(C_{f(k)}^{\mathsf{c}} \right) \right].$$

Since each $\{ \psi_n^{f(i)} \}_{n\in\mathbb{N}}$ is an a.s. α-verifier of $C_{f(i)}$, and $\mu \notin \cup_{i<k} C_{f(i)}$, it follows that $\mu^\infty \left[\limsup_{n\to\infty} (\psi_n^{f(i)})^{-1}(W) \right] = 0$ for $i < k$. Therefore,

$$\mu^\infty \left[\liminf_{n\to\infty} \lambda_n^{-1}(\mathscr{Q}_\mu) \right] \geq 1 - \mu^\infty \left[\limsup_{n\to\infty} \left(\psi_n^{f(k)} \right)^{-1} \left(C_{f(k)}^{\mathsf{c}} \right) \right].$$

Finally, since $\sum_{n=1}^\infty \mu^\infty \left[\left(\psi_n^{f(k)} \right)^{-1} \left(C_{f(k)}^{\mathsf{c}} \right) \right] \leq \alpha$, the Borel-Cantelli lemma yields:

$$\mu^\infty \left[\limsup_{n\to\infty} \left(\psi_n^{f(k)} \right)^{-1} \left(C_{f(k)}^{\mathsf{c}} \right) \right] = 0,$$

so $\mu^\infty \left[\liminf_{n\to\infty} \lambda_n^{-1}(\mathscr{Q}_\mu) \right] = 1$, as required.

2 entails 1. Immediate from the definitions. $\qquad\square$