THE CORROBORATION PARADOX (in press, *Synthese*)

Carl G. Wagner
Department of Mathematics
The University of Tennessee
Knoxville TN 37996-1320 USA
cwagner@tennessee.edu

Abstract. Evidentiary propositions E_1 and E_2, each p - positively relevant to some hypothesis H, are *mutually corroborating* if $p(H \mid E_1 \cap E_2) > p(H \mid E_i)$, $i = 1, 2$. Failures of such mutual corroboration are instances of what may be called the *corroboration paradox.* This paper assesses two rather different analyses of the corroboration paradox due, respectively, to John Pollock and Jonathan Cohen. Pollock invokes a particular embodiment of the principle of insufficient reason to argue that instances of the corroboration paradox are of negligible probability, and that it is therefore defeasibly reasonable to assume that items of evidence positively relevant to some hypothesis are mutually corroborating. Taking a different approach, Cohen seeks to identify supplementary conditions that are sufficient to ensure that such items of evidence will be mutually corroborating, and claims to have identified conditions which account for most cases of mutual corroboration. Combining a proposed common framework for the general study of paradoxes of positive relevance with a simulation experiment, we conclude that neither Pollock's nor Cohen's claims stand up to detailed scrutiny.

Keywords: corroboration, paradox, positive relevance, probability, simulation.

I am quite prepared to be told…"oh, that is an extreme case: it could never really happen!" Now I have observed that this answer is always given instantly, with perfect confidence, and without any examination of the proposed case. It must therefore rest on some general principle: the mental process being something like this—"I have formed a theory. This case contradicts my theory. Therefore, this is an extreme case, and would never occur in practice."

Rev. Charles L. Dodgson

1. The Corroboration Paradox. Let $E_1, \ldots, E_n$ be evidentiary propositions bearing on some hypothesis H, construing these propositions as subsets of some set Ω of possible states of the world. For all $I \subseteq [n] := \{1, \ldots, n\}$, let

(1.1) $E_I := \bigcap_{i \in I} E_i,$

where $E_\varnothing = \Omega$. Let $\mathcal{A}$ be the algebra generated by $E_1,\ldots,E_n$ and H, and let p be a probability measure on $\mathcal{A}$. If each E_i is $p-$ *positively relevant to* H ($p(H \mid E_i) > p(H)$), these propositions are said to be *mutually corroborating with respect to* H if

(1.2) $p(H \mid E_J) > p(H \mid E_I)$ whenever $[n] \supseteq J \supset I$.

The *corroboration paradox* is said to occur whenever a sequence of evidentiary propositions, each positively relevant to some hypothesis, fail to be mutually corroborating with respect to that hypothesis.[1] The fact that numerical examples of the corroboration paradox abound is unsurprising in view of the following theorem, which establishes that the numbers $p(H \mid E_I)$ may in general be utterly arbitrary.

Given $E_1,\ldots,E_n$ and H as above, and $I \subseteq [n]$, let

(1.3) $E_I^{\#} := E_I \cap (\bigcap_{i \in [n]-I} E_i^c)$.

In particular, $E_\varnothing^{\#} = \bigcap_{i \in [n]} E_i^c = (\bigcup_{i \in [n]} E_i)^c$, and $E_{[n]}^{\#} = E_{[n]} = \cap_{i \in [n]} E_i$.

The propositions $E_1,\ldots,E_n$ and H are said to be *qualitatively independent* (Rényi, 1970) if, for all $I \subseteq [n]$, $H \cap E_I^{\#} \neq \varnothing$ and $H^c \cap E_I^{\#} \neq \varnothing$.

THEOREM 1.1. If $E_1,\ldots,E_n$ and H are qualitatively independent, then for every family $\{c_I\}_{I \subseteq [n]}$ of real numbers belonging to the open interval $(0,1)$, there exists a probability measure p defined on the algebra $\mathcal{A}$ generated by $E_1,\ldots,E_n$ and H such that

(1.4) $p(H \mid E_I) = c_I$ for every $I \subseteq [n]$.

PROOF. See Appendix I. □

It follows from the above theorem that the patterns of support which the propositions E_i offer to H can be truly Byzantine. It is possible, for example, that each E_i is $p-$positively relevant to H, while two-at-a-time conjunctions of the E_i are $p-$negatively relevant to H, but three-at-a-time conjunctions are again $p-$positively relevant to H, and so on. This theorem explains *inter alia* why one may construct at will numerical examples of the corroboration paradox. Left unanswered, however, is the question of whether the possibility of encountering this paradox has any significance, either for cases of practical decision making or for Bayesian confirmation theory, which attempts to explicate confirmation as probabilistic positive relevance. In what follows, we examine two rather different analyses of this issue, one due to John Pollock and the other to Jonathan Cohen.[2] Pollock invokes the principle of insufficient reason for certain first and second order probabilities to argue that instances of the corroboration paradox are of negligible probability, and that it is therefore defeasibly reasonable to assume that items of evidence positively relevant to some hypothesis are mutually corroborating. Cohen, on the other hand, seeks to identify supplementary conditions that are sufficient to ensure that positively relevant items of evidence will be mutually corroborating. Pollock's approach is described below in Section 2, and Cohen's in Section 3. In Section 5 these approaches are evaluated within a proposed common framework for the general study of paradoxes of positive relevance.

2. Pollock on Corroboration.

Suppose that

(2.1) $$p(H) = a, \quad p(H \mid E_1) = r, \text{ and } \ p(H \mid E_2) = s, \text{ where } 0 < a, r, s < 1,$$

where $H, E_1,$ and E_2 are subsets of some set of possible worlds Ω. As is clear from Theorem 1.1, nothing whatsoever may be deduced from this information about the value of $p(H \mid E_1 \cap E_2)$. Nevertheless, invoking the principle of insufficient reason, John Pollock (2009) argues that if $a, r,$ and s are rational numbers, Ω is finite, and p is the uniform probability measure on subsets of Ω, then, for the uniform (second-order) probability P defined on subsets of triples (H, E_1, E_2) satisfying (2.1), the P-probability that E_1 and E_2 fail to be mutually corroborating can be made as near as we like to 0 for a sequence of finite sets Ω of increasing cardinality. In short, failures of mutual corroboration are, in the models constructed by Pollock, extreme cases. Pollock's argument is highly intricate, and relies on a computer algebra program that he wrote to establish certain key details. The following is a summary of his conclusions, expressed in standard mathematical terminology.

Let Ω be a finite set of possible states of the world, and let p denote the uniform probability measure on subsets of Ω, so that $p(E) := |E| / |\Omega|$ for all $E \subseteq \Omega$. Suppose that the cardinality of Ω is such that the family

(2.2) $$\mathcal{F}(\Omega, a, r, s) := \{(H, E_1, E_2) : H, E_1, E_2 \subseteq \Omega, \ p(H) = a, \ p(H \mid E_1) = r, \text{ and } \ p(H \mid E_2) = s\}$$

is nonempty.[3] Let P denote the uniform probability measure on subsets of $\mathcal{F}(\Omega, a, r, s)$, and define a random variable $X : \mathcal{F}(\Omega, a, r, s) \to [0,1]$ by $X(H, E_1, E_2) = p(H \mid E_1 \cap E_2)$. Let

(2.3) $$Y(r, s \mid a) := \frac{rs(1-a)}{a(1-r-s)+rs}.$$ [4]

THEOREM 2.1. (i) For all $K > 0$ there exists an $N > 0$ such that if $|\Omega|$ is an integer multiple of N, then $|\mathcal{F}(\Omega, a, r, s)| \geq K$. (ii) For all $\delta > 0$ and all $\varepsilon > 0$, there exists a K such that

(2.4) $$P(|X - Y(r,s|a)| \leq \delta) \geq 1 - \varepsilon \text{ whenever } |\mathcal{F}(\Omega, a, r, s)| \geq K.$$

PROOF. See Pollock (2009). □

It is easy to verify that $0 < Y(r,s|a) < 1$, and to establish the following theorem, which Pollock states without proof.

THEOREM 2.2. If $s > a$, then $Y(r,s|a) > r$ and if $r > a$, then $Y(r,s|a) > s$.

PROOF. See Appendix II. □

Although Pollock does not state the following theorem explicitly, it is implicit in Theorems 2.1 and 2.2.

THEOREM 2.3. The P-probability of the subset of $\mathcal{F}(\Omega, a, r, s)$ consisting of those triples (H, E_1, E_2) for which E_1 and E_2 are mutually p-corroborating with respect to H can be made as close to 1 as we wish on an infinite sequence of finite sets Ω of increasing cardinality.

PROOF. To make the P-probability in question ε-close to 1, let $\delta = (Y(r,s|a) - \max\{r,s\})/2$. Choose K sufficiently large so that (2.4) holds, and choose N such that if $|\Omega|$ is an integer multiple of N, then $|\mathcal{F}(\Omega, a, r, s)| \geq K$. □

Pollock's analysis, which seeks to reduce the corroboration problem to an exercise in enumerative combinatorics, has an undeniable charm. But the notorious instability of conclusions derived by the principle of insufficient reason, when one shifts from one set of possible states of the world to another equally plausible set of states, should incline one to caution regarding Pollock's conclusion that instances of the corroboration paradox are of negligible relative frequency. Indeed, as shown in Section 5 below, a differently structured application of this principle issues a very different verdict on the frequency with which one may expect to encounter this paradox.

3. Cohen on Corroboration. Jonathan Cohen (1977, 1980, 1986) has been a vigorous critic the application of classical probability (which he calls "mathematical," or "Pascalian" probability) to inductive logic, especially in the realms of legal and medical decision making. In its place, Cohen has proffered his own system of linearly ordered scores (which he calls "inductive," or "Baconian" probabilities, although they need not be numerical), and which are intended to record the degree to which a hypothesis has survived a sequence of attempts to falsify it. In an intriguing detour from his inductivist critique of classical probability, however, Cohen(1977, p. 101) asserts that there is a "demonstrably adequate analysis of corroboration…in terms of mathematical probability." Cohen's aim is not to argue that instances of the corroboration paradox are extreme cases. Rather, he aims to identify supplementary conditions that are sufficient to ensure that items of evidence positively relevant to some hypothesis are mutually corroborating.

Here is Cohen's analysis in the case of evidentiary propositions E_1 and E_2 satisfying

(3.1) $p(H \mid E_1) > p(H)$ and

(3.2) $p(H \mid E_2) > p(H)$.

The aim is to find supplementary conditions which, in the presence of (3.1) and (3.2), are sufficient to ensure that E_1 and E_2 are mutually corroborating with respect to H, i.e., that

(3.3) $p(H \mid E_1 \cap E_2) > p(H \mid E_1)$ and

(3.4) $p(H \mid E_1 \cap E_2) > p(H \mid E_2)$.

It is natural to expect, and straightforward to demonstrate, that (3.1) and (3.2), along with the conditional independence of E_1 and E_2, given H, and given H^c, imply (3.3) and (3.4). But as Cohen points out, this does not vouchsafe many cases of mutual corroboration, since conditional independence is a stringent condition that can only rarely be expected to obtain. Instead, Cohen calls our attention to the inspired generalizations of conditional independence[5] expressed by the inequalities

(3.5) $p(E_1 \mid E_2 \cap H) \geq p(E_1 \mid H)$ [equivalently, $p(E_2 \mid E_1 \cap H) \geq p(E_2 \mid H)$] and

(3.6) $p(E_1 \mid E_2 \cap H^c) \leq p(E_1 \mid H^c)$ [equivalently, $p(E_2 \mid E_1 \cap H^c) \leq p(E_2 \mid H^c)$],

and proves, as a consequence of the following theorem, that (3.1) and (3.2), supplemented by (3.5) and (3.6), are sufficient to ensure that E_1 and E_2 are mutually corroborating with respect to H:

THEOREM 3.1. Conditions (3.1), (3.5), and (3.6) together imply condition (3.4). Conditions (3.2), (3.5), and (3.6) together imply condition (3.3).

PROOF. Since Cohen's proof of this theorem runs to three pages, the following short proof may be of interest: It is easy to verify that if q is any probability measure on $\mathcal{A}$, then $q(H \mid E) > q(H)$ if and only if $q(E \mid H) > q(E \mid H^c)$.[6] Letting $q = p$ and $E = E_1$, it follows that (3.1) is equivalent to $p(E_1 \mid H) > p(E_1 \mid H^c)$. Letting $q(\cdot) = p(\cdot \mid E_2)$ and $E = E_1$, it follows that (3.4) is equivalent to $p(E_1 \mid E_2 \cap H) > p(E_1 \mid E_2 \cap H^c)$. But by (3.5) and (3.6), (3.7) $$p(E_1 \mid E_2 \cap H) \geq p(E_1 \mid H) > p(E_1 \mid H^c) \geq p(E_1 \mid E_2 \cap H^c),$$

which establishes (3.4). Interchanging E_1 and E_2 in the preceding argument, replacing (3.1) with (3.2), and (3.4) with (3.3), and applying the bracketed equivalents of (3.5) and (3.6), establishes (3.3). □

Remark 3.1. In proving that (3.1), (3.5), and (3.6) imply (3.4), we did not assume (3.2). Since we have restricted the term "corroboration" to apply to evidentiary propositions assumed to be positively relevant to some hypothesis, condition (3.4) alone does not assert that E_1 $p-$corroborates E_2 with respect to H. Rather, (3.4) simply states that E_1 is conditionally $p-$positively relevant to H, given E_2, which we might paraphrase by saying that E_1 $p-$*reinforces* E_2 *with respect to* H. Similar remarks apply to condition (3.5) when (3.1) is not assumed.

Remark 3.2. There are straightforward generalizations of conditions (3.5) and (3.6) which, along with the $p-$positive relevance of $E_1, E_2, \ldots$, and E_n to H, ensure that the propositions E_i are mutually corroborating with respect to H. See Appendix III.

The heuristics employed by Cohen to discover conditions (3.5) and (3.6) are interesting to examine. For concreteness, he considers the case in which H is a proposition material to some legal proceeding, E_1 is the testimony of witness 1 that H is true, E_2 is the testimony of witness 2 that H is true, and (3.1) and (3.2) are assumed to hold. Suppose now that (3.6) failed to hold, so that

(3.8) $$p(E_1 \mid E_2 \cap H^c) > p(E_1 \mid H^c) \quad \text{and} \quad p(E_2 \mid E_1 \cap H^c) > p(E_2 \mid H^c).$$

Then each witness is more inclined to give false testimony when the other does, and Cohen (1977, p.101) asserts that "if one witness is to corroborate another…, one witness must not be more inclined to give false testimony when the other does." Similarly, suppose that (3.5) failed, so that

(3.9) $$p(E_1 \mid E_2 \cap H) < p(E_1 \mid H) \quad \text{and} \quad p(E_2 \mid E_1 \cap H) < p(E_2 \mid H) .$$

Then each witness is less inclined to agree with the other when the latter's testimony is true, and Cohen (1977, p. 102) asserts that "if one witness is to corroborate another, the former's inclination to give true testimony…must not be reduced when the latter's testimony is true." Taken literally, Cohen appears to be asserting that conditions (3.6) and (3.5) are *necessary* for corroboration,[7] but a subsequent remark of his, to be discussed immediately below, indicates that he did not mean to make these strong (and incorrect) claims. Rather, what he seemed to have in mind was that a juror who believed (3.8) or (3.9) might reasonably find the agreement of the witnesses in asserting the truth of H to be *suspect*. In the case of (3.8) the suspicion might be that there was a conspiracy to deceive. In the case of (3.9) such agreement might be suspected to have come about due to the falsehood of H.

That Cohen did not mean to assert the absolute necessity of conditions (3.5) and (3.6) for mutual corroboration is apparent from his claim in Cohen (1980, p. 51) that such corroboration "will not normally" take place unless (3.5) and (3.6) hold. On the other hand, he appears to be making the only slightly less strong claim that, among cases in which positively relevant items of evidence are mutually corroborating, it will only rarely be the case that (3.5) or (3.6) fails to hold. This is clearly a second-order probability claim, though not supported by any substantive arguments. The need for such arguments is underscored in the next section, where it is shown that, for at least one possible second-order probability, in only around 39 per cent of the cases of positively relevant, mutually corroborating items of evidence is the corroboration accounted for by (3.5) and (3.6).

4. A Framework for Assessing Paradoxes in the Theory of Probability. Let $\mathcal{A}$ be a Boolean algebra of propositions, and let $\Pi_{\mathcal{A}}$ be the set of all probability measures on $\mathcal{A}$, with a generic member of $\Pi_{\mathcal{A}}$ denoted by p. In what follows, the symbols $\mathcal{C}$, $\mathcal{D}$, and $\mathcal{S}$ denote subsets of $\Pi_{\mathcal{A}}$ or, alternatively, predicates on the domain $\Pi_{\mathcal{A}}$; we write $\mathcal{C}(p)$ whenever $p \in \mathcal{C}$, and similarly for $\mathcal{D}$ and $\mathcal{S}$. In general, probability paradoxes involve a condition $\mathcal{C}(p)$, which naïve intuition (mistakenly) suggests should entail some desirable condition $\mathcal{D}(p)$, i.e., that $\mathcal{C} \subseteq \mathcal{D}$. Upon realizing that $\mathcal{C}(p)$ does not imply $\mathcal{D}(p)$, one is naturally inclined to try to identify some supplementary condition $\mathcal{S}(p)$ which, in conjunction with $\mathcal{C}(p)$, does entail $\mathcal{D}(p)$, i.e., such that $\mathcal{C} \cap \mathcal{S} \subseteq \mathcal{D}$. Such supplementary conditions have several important functions:

(i) In experimental situations, one can often "design in" the condition $\mathcal{S}$(p), thereby ensuring that $\mathcal{D}(p)$ will hold whenever $\mathcal{C}(p)$ does. Suppose, for example, that one wishes to test the efficacy of a treatment for a certain disease using experimental subjects from two different hospitals. Let Ω denote the set of all experimental subjects, T the subset of Ω consisting of all those subjects receiving the treatment, R the subset consisting of all those subjects, treated or untreated, who recover from the disease after a certain time, and H_1 and H_2 the set of subjects at hospitals 1 and 2, respectively. Let $p(E) := |E|/|\Omega|$ for all $E \subseteq \Omega$. As is well-known, it is possible that $p(R|T \cap H_i) > p(R|T^c \cap H_i)$ for $i = 1, 2$, yet $p(R|T) < p(R|T^c)$. But this possibility, which is one example of a general phenomenon known as *Simpson's paradox*[8], can easily be avoided if one simply administers the treatment to the same fraction of subjects in H_1 as in H_2. [9]

(ii) When encountering cases where $\mathcal{C}(p)$ does not entail $\mathcal{D}(p)$, one knows immediately that $\mathcal{S}(p)$ must fail to hold and, in identifying the precise nature of this failure, one gains useful insights into the phenomenon under consideration. Consider, for example, the following simplified version of the famous Berkeley admissions case (Bickel et al, 1975), which also involves an example of Simpson's paradox. Here, Ω denotes the set of all applicants for admission to the graduate programs of a certain university, with H denoting the set of applicants to humanities departments, S the set of applicants to science departments (where $\{H, S\}$ partitions Ω), F the set of female applicants, M the set of male applicants, and A the set of applicants accepted by the department to which they applied. Again, $p(E) := |E|/|\Omega|$. Suppose that that $\mathcal{C}(p) \equiv [\, p(A|F \cap H) \geq p(A|M \cap H)$ and $p(A|F \cap H) \geq p(A|M \cap H)]$, yet $p(A|F) < p(A|M)$. Since $\mathcal{S}(p) \equiv [p(A|H) = p(A|S)$ or $p(F|H) = p(F|S)]$ would guarantee that $p(A|F) \geq p(A|F)$,[10] we know both that $p(A|H) \neq p(A|S)$ and that $p(F|H) \neq p(F|S)$. Indeed, closer examination of the data reveals that $P(F|H) > P(F|S)$ and $p(A|H) < p(A|S)$, and to such an extent as to explain the lower acceptance rate for female applicants overall.

(iii) The foregoing examples typically involve empirical probabilities, i.e., relative frequencies ascertained by design or observation. But there is also a role for $\mathcal{S}(p)-$type conditions in the formulation of subjective assessments of probability, especially in the case of qualitative probability assessments (expressed by relations such as "is more probable than," or "is at least as probable as"). For the predicates $\mathcal{C}$, $\mathcal{S}$, and $\mathcal{D}$ typically involve probability inequalities, and can thus be interpreted in subjective contexts as qualitative probability judgments. $\mathcal{S}(p)-$type conditions then function in the following way: Having made the qualitative probability judgment $\mathcal{C}(p)$, and contemplating whether to embrace the judgment $\mathcal{D}(p)$, one can approach this decision indirectly by considering whether it is reasonable to judge that $\mathcal{S}(p)$. When $\mathcal{S}$ is more salient than $\mathcal{D}$, this may be a useful strategy for the construction of qualitative probability judgments. Unless it should be the case that, in the presence of $\mathcal{C}(p)$, $\mathcal{S}(p)$ is necessary, as well as sufficient, for $\mathcal{D}(p)$, rejecting $\mathcal{S}(p)$ will of course still leave the decision about whether to embrace $\mathcal{D}(p)$ unresolved.

5. Probable Probabilities.

We are now prepared to deal with assertions to the effect that a given paradox has negligible chance of occurring, or that a given supplementary condition is responsible for most cases in which both $\mathcal{C}(p)$ and $\mathcal{D}(p)$ hold. These are, as noted earlier, second-order probability claims, and demand an explicit representation in terms of such probabilities. To that end, let us suppose that the set $\Pi_{\mathcal{A}}$ of all probability measures on $\mathcal{A}$ is equipped with a sigma algebra Σ, and that P is a probability measure on Σ. Suppose further that $\mathcal{C}$, $\mathcal{S}$, and $\mathcal{D}$ are all members of Σ, and that P assigns events $\mathcal{C}$ and $\mathcal{D}$ nonzero probability. Then

(5.1) $$P(\mathcal{D}^{c}|\mathcal{C}) = 1 - P(\mathcal{D}|\mathcal{C})$$

measures what we shall call the *prevalence* of the paradox in question, relative to P, (5.2) $P(\mathcal{S}|\mathcal{C})$

measures what we shall call the *incidence of* $\mathcal{S}$ *in* $\mathcal{C}$, relative to P, and

(5.3) $P(\mathcal{S}|\mathcal{D}\cap\mathcal{C})$

measures what we shall call the $\mathcal{S}$*-provenance of* $\mathcal{D}$ *in* $\mathcal{C}$, relative to P. Since $\mathcal{C} \cap \mathcal{S} \subseteq \mathcal{D}$,

(5.4) $P(\mathcal{S}|\mathcal{C}) = P(\mathcal{D}|\mathcal{C})P(\mathcal{S}|\mathcal{D}\cap\mathcal{C})$,

i.e., "incidence = (1 – prevalence)$\times$provenance."

Let us now apply the foregoing framework to the corroboration paradox, with $\mathcal{A}$ being the algebra of propositions generated by E_1, E_2, and H,

(5.5) $\mathcal{C} = \{p \in \Pi_{\mathcal{A}}:\ p(H\,|\,E_1) > p(H)$ and $p(H\,|\,E_2) > p(H)\}$,

(5.6) $\mathcal{D} = \{p \in \Pi_{\mathcal{A}}:\ p(H\,|\,E_1 \cap E_2) > p(H\,|\,E_1)$ and $p(H\,|\,E_1 \cap E_2) > p(H\,|\,E_2)\}$, and

(5.7) $\mathcal{S} = \{p \in \Pi_{\mathcal{A}}:\ p(E_1\,|\,E_2 \cap H) \geq p(E_1\,|\,H)$ and $p(E_1\,|\,E_2 \cap H^c) \leq p(E_1\,|\,H^c)\}$.

In order to define a sigma algebra Σ of subsets of $\Pi_{\mathcal{A}}$ and a probability measure P on Σ, we identify each $p \in \Pi_{\mathcal{A}}$ with the 7-tuple $(p_1, p_2, ..., p_7) =$ $(p(HE_1E_2), p(HE_1E_2^c), p(HE_1^cE_2), p(HE_1^cE_2^c), p(H^cE_1E_2), p(H^cE_1E_2^c), p(H^cE_1^cE_2))$, where set intersection is indicated by concatenation. This results in an identification of $\Pi_{\mathcal{A}}$ with the simplex $\Pi_7 = \{(p_1, p_2, ..., p_7): p_i \geq 0$ and $\sum_{i=1}^{7} p_i \leq 1\}$. As an illustration, let us equip the sigma algebra of Lebesgue measurable subsets of Π_7 with the uniform probability (i.e., normalized Lebesgue) measure P. Since $\mathbb{P}_7$ has Lebesgue measure $\mu(\Pi_7) = 1/7!$, we set $P(E) = 7!\mu(E)$ for every Lebesgue measurable $E \subseteq \Pi_7$. The second-order probabilities $P(\mathcal{D}|\mathcal{C})$, $P(\mathcal{S}|\mathcal{C})$, and $P(\mathcal{S}|\mathcal{D}\cap\mathcal{C})$ are now well-defined, but their exact evaluation requires the computation of some complicated multiple integrals. Fortunately, one can approximate these integrals using Monte Carlo methods (see Appendix IV for details). A sample of one million probability measures p on $\mathcal{A}$, selected uniformly at random from $\Pi_{\mathcal{A}}$ (i.e., from Π_7) yielded the approximations

(5.7) $P(\mathcal{D}^c|\mathcal{C}) \approx 0.37$,

(5.8) $P(\mathcal{S}|\mathcal{C}) \approx 0.25$, and

(5.9) $P(\mathcal{S}|\mathcal{D}\cap\mathcal{C}) \approx 0.39$.

I should emphasize that in employing the uniform probability measure P above, I am in no way asserting that P is the "right" second-order probability for analyzing the corroboration paradox. Rather, P is introduced as a test case, and serves a cautionary role, highlighting the fact that the claims of Pollock and Cohen are overly broad and insufficiently supported.

The result $P(\mathcal{D}^c|\mathcal{C}) \approx 0.37$ shows that a differently structured application of the principle of insufficient reason yields an estimate of the prevalence of the corroboration paradox quite different from Pollock's. This is no surprise, since the results of applying this principle are notoriously unstable when one shifts from one set of possible states of the world to another equally plausible set of states. Absent a (highly unlikely) demonstration that Pollock's particular application of this principle is the uniquely rational way to assess the corroboration paradox, the result (5.7) simply devastates his claim that it is defeasibly reasonable to assume that two items of evidence, each positively relevant to some hypothesis, are mutually corroborating with respect to that hypothesis.

What light do these results shed on Cohen's claim that E_1 and E_2 will "not normally" be mutually corroborating with respect to H unless the supplementary conditions (3.5) and (3.6) hold? Cohen is clearly asserting that the $\mathcal{S}$-provenance of $\mathcal{D}$ in $\mathcal{C}$ is high, i.e., that property $\mathcal{S}$, as expressed in (3.5) and (3.6), accounts for the vast majority of cases in which evidentiary propositions E_1 and E_2, each positively relevant to hypothesis H, are mutually corroborating. But (5.9) indicates that, for the uniform probability measure P, property $\mathcal{S}$ accounts for only around 39% of such cases. This is far from a refutation of Cohen's claim. But it does underline the fact that anyone seeking to justify that claim must defend (as Cohen has not) a second-order probability measure, necessarily different from P, for which the $\mathcal{S}$-provenance of $\mathcal{D}$ in $\mathcal{C}$ is high. This may well involve distinct analyses of corroboration by means of case studies in law, medicine, and the various sciences, and while guided by the spirit of the framework introduced in Section 4 above, such analyses need not match the level of formality posited by that framework.[11] Even if such a program fails to produce convincing evidence for the high $\mathcal{S}$-provenance of $\mathcal{D}$ in $\mathcal{C}$, however, Cohen's identification of the $\mathcal{S}$ – conditions (3.5) and (3.6) furnishes an invaluable perspective on the phenomenon of corroboration.

Appendix I

THEOREM 1.1. If $E_1,...,E_n$ and H are qualitatively independent, then for every family $\{c_I\}_{I\subseteq[n]}$ of real numbers belonging to the open interval $(0,1)$, there exists a probability measure p defined on the algebra $\mathcal{A}$ generated by $E_1,...,E_n$ and H such that

(I.1) $p(H\mid E_I)=c_I$ for every $I\subseteq[n]$.

PROOF. For every $I\subseteq[n]$, select a number x_I according to the following recursive scheme: (i) Let $x_\varnothing=1$. (ii) The numbers x_I having been chosen for all I of cardinality r, choose numbers $x_J>0$ for every J of cardinality $r+1$ so that, for every I of cardinality r,

(I.2) $$c_I x_I-\sum_{J\supset I\,\&|J-I|=1} c_J x_j\geq 0,$$ and

(I.3) $$(1-c_I x_J)-\sum_{J\supset I\,\&|J-I|=1}(1-c_j)x_j\geq 0.$$

Now define a function p on all propositions of the form $H\cap E_I^{\#}$ and $H^c\cap E_I^{\#}$ by

(I.4) $$p(H\cap E_I^{\#}):=\sum_{J\supseteq I}(-1)^{|J-I|}c_J x_j,$$ and

(I.5) $$p(H^c\cap E_I^{\#}):=\sum_{J\supseteq I}(-1)^{|J-I|}(1-c_J)x_J,$$

extending p to unions of these mutually exclusive propositions in the obvious way so that p is additive on $\mathcal{A}$. To show that p is actually a probability measure on $\mathcal{A}$, it then suffices to show that the quantities (I.4) and (I.5) are all nonnegative, and that $p(\Omega)=1$.

From (I.4) we have

(I.6) $$p(H\cap E_I^{\#})=\{c_I x_I-\sum_{\substack{J\supset I\\|J-I|=1}} c_J x_J\}+\sum_{\substack{J\supset I\\|J-I|\geq 2}}(-1)^{|J-I|}c_J x_J.$$

The bracketed term on the right hand side of equation (I.6) is nonnegative by (I.2), and the remaining sum clearly dominates

(I.7) $$\sum_{\substack{J\supset I\\|J-I|\equiv 0(\mathrm{mod}\,2)}}\{c_J x_J-\sum_{\substack{K\supset J\\|K-J|=1}} c_K x_K\},$$

each term of which is again nonnegative by (I.2). The proof that the numbers $p(H^c \cap E_I^\#)$ are nonnegative is similar. Next, note that for every $I \subseteq [n]$, $(H \cap E_I) = \bigcup_{K \supseteq I} (H \cap E_K^\#)$, whence

(I.8) $$p(H \cap E_I) = \sum_{K \supseteq I} p(H \cap E_K^\#) = \sum_{K \supseteq I} \sum_{J \supseteq K} (-1)^{|J-K|} c_J x_J = \sum_{J \supseteq I} c_J x_J \sum_{J \supseteq K \supseteq I} (-1)^{|J-K|} = c_I x_I ,$$

since, if $J \supset I$, with $|J - I| = s > 0$,

(I.9) $$\sum_{J \supseteq K \supseteq I} (-1)^{|J-K)} = \sum_{i=0}^{s} (-1)^i \binom{s}{i} = (1-1)^s = 0.$$

Similarly, for every $I \subseteq [n]$,

(I.10) $$p(H^c \cap E_I^\#) = (1 - c_I) x_I ,$$

and so

(I.11) $$p(E_I) = p(H \cap E_I) + p(H^c \cap E_I) = x_I .$$

In particular, $p(\Omega) = p(E_\varnothing) = x_\varnothing = 1$, and so p is a probability measure. Finally, from (I.8) and (I.11), it follows that $p(H \mid E_I) = c_I$ for all $I \subseteq [n]$. □

Note that the above method yields infinitely many probability measures p on $\mathcal{A}$ satisfying (I.1), and that the set of all p satisfying (I.1) is convex, i.e., closed under weighted arithmetic averaging.

Appendix II

THEOREM 2.2. Let $Y(r,s \mid a) = \dfrac{rs(1-a)}{a(1-r-s)-rs}$, where $0 < a, r, s < 1$. If $s > a$, then $Y(r,s \mid a) > r$ and if $r > a$, then $Y(r,s \mid a) > s$.

PROOF. Since $Y(r,s \mid a)$ is symmetric in r and s, it suffices to prove the first of these assertions. Note first that $a(1-r-s)+rs = [(1-r)a + rs] - as \geq \min\{a,s\} - as > 0$, so $Y(r,s \mid a)$ is well-defined and positive. From the fact that $sa(1-r) < a(1-r)$ it follows that $rs(1-a) < a(1-r-s)+rs$, and so $Y(r,s \mid a) < 1$. Suppose that $s > a$. T $s(1-r) - as > a(1-r) - as$ $\Leftrightarrow$ $s(1-a) > a(1-r-s)+rs$ $\Leftrightarrow$ $\dfrac{s(1-a)}{a(1-r-s)+rs} > 1$ $\Leftrightarrow$ $\dfrac{rs(1-a)}{a(1-r-s)+rs} > r$. □

Note that the above theorem remains true when each of the inequalities $>$ is replaced by $<$.

Appendix III

GENERALIZED THEOREM 3.1. Let $E_1, ..., E_n$ be evidentiary propositions bearing on the hypothesis H, with E_I defined for all $I \subseteq [n]$ by (1.1). If

(II.1) $p(H \mid E_i) > p(H)$ for all $i \in [n]$,

(II.2) $p(E_i \mid E_I \cap H) \geq p(E_i \mid H)$ for all $I \subset [n]$ and all $i \in [n] - I$, and

(II.3) $p(E_i \mid E_I \cap H^c) \leq p(E_i \mid H^c)$ for all $I \subset [n]$ and all $i \in [n] - I$,

then

(II.4) $p(H \mid E_J) > p(H \mid E_I)$ whenever $J \supset I$.

PROOF. Let $E_1 = E_i$ and $E_2 = E_I$ in Theorem 3.1. This yields $p(H \mid E_i \cap E_I) > p(H \mid E_I)$, from which (II.4) follows by induction on $|J - I|$. □

Appendix IV

One simulates the uniform distribution on $\Pi_n := \{(p_1, ..., p_n) : 0 \leq p_i \leq 1 \text{ and } p_1 + \cdots + p_n \leq 1\}$ by using the following theorem:

THEOREM IV.1. If $U_1, ..., U_{n+1}$ are uniformly distributed on the open interval $(0,1)$, $Y_i := -\ln U_i$, and $S := Y_1 + \cdots + Y_{n+1}$, then the random vector $(X_1, ..., X_n)$, with $X_i := Y_i / S$, is uniformly distributed on Π_n.

PROOF: It is easy to show by induction on n that the Lebesgue measure $\mu(\Pi_n) = 1/n!$. So to show that X is uniform on Π_n, it suffices to show that the density f_X of X takes the form $f_X(x_1, ..., x_n) = n!$ if $(x_1, ..., x_n) \in \Pi_n$ and $f_X(x_1, ..., x_n) = 0$ otherwise. Let $g : \mathbb{R}^{n+1} \to \mathbb{R}$ be a bounded function. Then

$$\text{(IV.1)} \qquad Eg(X_1, ..., X_n, S) = Eg\left(\frac{Y_1}{S}, ..., \frac{Y_n}{S}, S\right) = \int_0^\infty \cdots \int_0^\infty g\left(\frac{y_1}{s}, ..., \frac{y_n}{s}, s\right) e^{-y_1} \cdots e^{-y_n} e^{-y_{n+1}} dy_1 \cdots dy_{n+1},$$

since each Y_i is exponentially distributed, with parameter 1. Under the change of variables $x_1 = \frac{y_1}{s}, ..., x_n = \frac{y_n}{s}$, where $s = y_1 + \cdots + y_{n+1}$, we have $y_1 = sx_1, ..., y_n = sx_n$, and $y_{n+1} = s - \sum_{i=1}^{n} sx_i$, from which it follows that the Jacobian

$$J = \det \begin{bmatrix} s & 0 & \cdots & 0 & x_1 \\ 0 & s & 0 & \cdots & x_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdots & 0 & s & x_n \\ -s & -s & \cdots & -s & 1-\sum_{i=1}^{n} x_i \end{bmatrix} = \det \begin{bmatrix} s & 0 & \cdots & 0 & x_1 \\ 0 & s & \cdots & 0 & x_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & s & x_n \\ 0 & \cdot & \cdot & 0 & 1 \end{bmatrix} = s^n.$$

Hence the integral in (IV.1) is equal to

(IV.2) $$\int_{\Pi_n} \int_0^{\infty} g(x_1, ..., x_n, s) e^{-s} s^n dx_1 \cdots dx_n ds.$$

Now set $g(x_1, ..., x_n, s) = h(x_1, ..., x_n)$. Then

$$Eh(X_1, ..., X_n) = \int_{\Pi_n} \int_0^{\infty} h(x_1, ..., x_n) e^{-s} s^n dx_1 \cdots dx_n ds = \int_{\Pi_n} h(x_1, ..., x_n) dx_1 \cdots dx_n \int_0^{\infty} e^{-s} s^n ds$$

$= \int_{\Pi_n} h(x_1, ..., x_n) n! dx_1 \cdots dx_n$, by Euler's identity. Given any measurable set $A \subseteq \mathbb{R}^n$, setting $h = \chi_A$, the characteristic function of A, completes the proof that X is uniform on Π_n. □

Remark 1. I am indebted to my friend and colleague Jan Rosinski for his crucial assistance in the formulation and proof of the above procedure for simulating the uniform distribution on Π_n.

Remark 2. If exponential random variables $Y_1, ..., Y_{n+1}$ with parameter 1 can be generated directly from a computer program, the initial step in the above procedure can be omitted.

Notes

1. In this paper, the term "paradox" denotes an *apparent*, not actual, contradiction. The corroboration paradox is just one example of the many paradoxes of positive relevance. In most cases the paradoxical air of such examples derives from mistakenly construing positive relevance

as an attenuated sort of implication. Unlike implication, however, the positive relevance relation is symmetric, and non-transitive. Indeed, just about all of one's naïve expectations regarding positive relevance turn out to be false. Chung (1942) provides a comprehensive catalogue of examples.

2. Both Pollock and Cohen restrict consideration to the case of two evidentiary propositions. Cohen's analysis is, as we shall see, easily extended to any finite number of such propositions. There appear, however, to be formidable roadblocks to a similar extension of Pollock's analysis.

3. Suppose that the rationals $a, r,$ and s are expressed as fractions in lowest terms, with $a = \alpha^* / \alpha,\ r = \rho^* / \rho,$ and $s = \sigma^* / \sigma.$ Since $\alpha | H | = \alpha^* | \Omega |$, it must be the case that $|\Omega|$ is an integer multiple of α, say $|\Omega| = m\alpha,$ whence $|H| = m\alpha^*$. Since $\rho | H \cap E_1 | = \rho^* | H |$ and $\sigma | H \cap E_2 | = \sigma^* | H |,$ it must be the case that $|H|$ is divisible by both ρ and σ, and hence that $m\alpha^*$ is an integer multiple of the least common multiple of ρ and σ.

4. It is interesting to note that $Y(r, s | 1/2) = rs / [rs + (1-r)(1-s)]$ agrees with the result of applying Dempster's rule of composition for belief functions (Shafer, 1976). See Pollock (2009) for an elaboration.

5. The conditional independence of E_1 and E_2, given H, and given H^c, are just the respective equality cases of the inequalities (3.5) and (3.6).

6. One simply shows that each of these inequalities is equivalent to the inequality $q(E \cap H)q(E^c \cap H^c) > q(E^c \cap H)q(E \cap H^c).$ The foregoing inequality expresses the "determinant test" for the relation of positive relevance between E and H. Indeed, with $a = q(E \cap H),\ b = q(E \cap H^c),\ c = q(E^c \cap H),$ and $d = q(E^c \cap H^c),$ E and H are, respectively, $q-$positively relevant to each other, $q-$negatively relevant to each other, or $q-$independent according as the determinant of the matrix $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is positive, negative, or equal to zero.

7. Schlesinger (1988) is among those misled on that score. See Wagner (1991).

8. See Simpson (1951) and Blyth (1972)

9. Since $\{H_1, H_2\}$ is a partition of Ω, it follows from $P(T | H_1) = P(T | H_2)$ that $P(H_i | T) = P(H_i | T^c),\ i = 1, 2.$ Hence, if (i) $P(R | H_i \cap T) > P(R | H_i \cap T^c)$, $i = 1, 2,$ then

$$P(R | T) = P(R \cap H_1 | T) + P(R \cap H_2 | T) = P(H_1 | T)P(R | H_1 \cap T) + P(H_2 | T)P(R | H_2 \cap T)$$

$> P(H_1 | T^c)P(R | H_1 \cap T^c) + P(H_2 | T^c)P(R | H_2 \cap T^c) = P(R | T^c).$ It is also true (and left as an exercise to show) that (i), in conjunction with the supplementary condition $P(R | H_1) = P(R | H_2)$, implies that $P(R | T) > P(R | T^c)$. But this is of no interest here, since we cannot "design in" condition (ii).

10. The proofs are similar to the proof in Note 9 *supra.*

11. It is interesting, and ironic, that a formal and experimental (to the extent that simulation counts as experimentation) study of corroboration should wind up reminding us of the value of good old-fashioned case studies. It is also conceivable that such studies may identify clusters of cases in which the *prevalence* of the corroboration paradox is negligible. This would be of no help to Pollock, however, since his analysis is meant to be completely general in application, and is unalterably wedded to particular embodiments of the principle of insufficient reason, for both first and second order probabilities.

References

P.J. Bickel, E.A. Hammel, and J. W. O'Connell (1975). Sex bias in graduate admissions: data from Berkeley, *Science* 187, 398-404.

C.R. Blyth (1972). On Simpson's paradox and the sure thing principle, *Journal of the American Statistical Association* 67, 364-366.

K.-L. Chung (1942). On mutually favorable events, *Annals of Mathematical Statistics* 13, 338-349.

L. J. Cohen (1977). *The Probable and the Provable*, Clarendon Press, Oxford.

L. J. Cohen (1980). Bayesianism versus Baconianism in the evaluation of medical diagnoses, *British Journal for the Philosophy of Science* 31, 45-62.

J. Pollock (1990). *Nomic Probability and the Foundations of Induction*, Oxford University Press, New York.

J. Pollock (2009). Probable probabilities, available online at: http://oscarhome.soc-sci.arizona.edu/ftp/PAPERS/Probable Probabilities with proofs.pdf

A. Rényi (1970). *Foundations of Probability,* Holden Day, San Francisco.

G. Schlesinger (1988). Why a tale twice told is more likely to take hold, *Philosophical Studies* 54, 141-152.

G. Shafer (1976). *A Mathematical Theory of Evidence*, Princeton University Press, Princeton.

E.H. Simpson (1951). The interpretation of interaction in contingency tables, *Journal of the Royal Statistical Society* (Series B) 13, 238-241.

C. Wagner (1991). Corroboration and conditional positive relevance, *Philosophical Studies* 61, 295-300.