PATRICK MAHER

PROBABILITIES FOR MULTIPLE PROPERTIES: THE MODELS OF HESSE AND CARNAP AND KEMENY

ABSTRACT. In 1959 Carnap published a probability model that was meant to allow for reasoning by analogy involving two independent properties. Maher (2000) derived a generalized version of this model axiomatically and defended the model's adequacy. It is thus natural to now consider how the model might be extended to the case of more than two properties. A simple extension was published by Hesse (1964); this paper argues that it is inadequate. A more sophisticated one was developed jointly by Carnap and Kemeny in the early 1950s but never published; this paper gives the first published description of Carnap and Kemeny's model and argues that it too is inadequate. Since no other way of extending the two-property model is currently known, the conclusion of this paper is that a satisfactory extension to multiple properties requires some new approach.

1. INTRODUCTION

Argument by analogy is a generally accepted form of inductive reasoning and many think that inductive reasoning can be represented using the probability calculus. From these facts one might expect that there would be accepted probability models that can represent inference by analogy, but no such model exists. This paper will explore some of the obstacles to creating such a model.

I will begin by describing the domain of the probability models with which I will be concerned. Let F_1^i, $i = 1, \dots, n$, be logically independent properties and let F_2^i be the negation of F_1^i. Let a population of individuals be also given. Let an atomic proposition be a proposition that ascribes one of the F_l^i to one of the individuals. Let $\mathcal{X}$ be the algebra generated by the atomic propositions, that is, the smallest set of propositions that contains the atomic propositions and is closed under conjunction and negation. The probability models with which I will be concerned are sets of probability functions defined on $\mathcal{X}$.

It is most useful to find probability functions that are appropriate when there is no relevant evidence, since by conditionalization one can then obtain probability functions that are appropriate given any specified evidence. So let R denote the class of probability functions on $\mathcal{X}$ that represent

Erkenntnis **55:** 183–216, 2001.

rationally permissible degrees of certainty when there is no relevant evidence. (Subjectivists may replace 'rationally permissible' by 'acceptable to me' or 'acceptable to many people'.) A probability model will be useful if it is either a superset or a subset of R; membership in the model will then be a necessary or sufficient condition, respectively, for a probability function to be rationally permissible.

The contents of R will in general depend on the interpretation of the F^i_l. For example, some standard forms of argument by analogy will be inappropriate if we use properties like Grue. (I use capitalized predicates to denote properties.) I will assume that what we want is a model that is useful when the properties are fairly normal ones. For example, we might take F^1_1 to be Swan, F^2_1 to be Australian, and F^3_1 to be White.

2. TERMINOLOGY AND NOTATION

In this section I will introduce some terminology and notation that will be used throughout the paper. I will use $F^{i_1 \ldots i_k}_{l_1 \ldots l_k}$ to denote the property of having all of $F^{i_1}_{l_1}, \ldots, F^{i_k}_{l_k}$. So in my swan example, F^{12}_{12} is Non-Australian Swan and F^{123}_{211} is White Australian Non-Swan.

A *family of properties* is a set of properties that are pairwise exclusive and jointly exhaustive. For any distinct $i_1, \ldots, i_k \in \{1, \ldots, n\}$ I will use $\mathcal{F}^{i_1 \ldots i_k}$ to denote the family of properties

$$\{F^{i_1 \ldots i_k}_{l_1 \ldots l_k} : l_1, \ldots, l_k \in \{1, 2\}\}.$$

For example, $\mathcal{F}^1 = \{F^1_1, F^1_2\}$ and $\mathcal{F}^{12} = \{F^{12}_{11}, F^{12}_{12}, F^{12}_{21}, F^{12}_{22}\}$.

For any property ϕ and individual a I will use ϕa to denote that a has ϕ. Also, thinking of propositions as sets of states or models, the conjunction of propositions A and B will be represented by the set intersection $A \cap B$.

A *sample* is a finite subset of the set of individuals. A *sample proposition* with respect to family of properties Φ is a proposition that ascribes a property from Φ to each member of some sample. For example, $F^1_1 a \cap F^1_2 b \cap F^1_2 c$ is a sample proposition with respect to $\mathcal{F}^1$. It is convenient to allow that the empty set is a sample and the necessarily true proposition is a sample proposition for the empty set with respect to any family.

DEFINITION 1. A family of properties Φ is a λ*-family relative to probability function* p iff there exists $\lambda > 0$ and for each $\phi \in \Phi$ there exists $\gamma_\phi \in (0, 1)$ such that the following holds: If E is a sample proposition with respect to Φ involving s individuals and if s_ϕ is the number of individuals to which E ascribes property ϕ then for any individual a not involved in E,

$$p(\phi a|E) = \frac{s_\phi + \lambda\gamma_\phi}{s + \lambda}.$$

The properties ϕ with which I will be concerned have the form $F^{i_1 \ldots i_k}_{l_1 \ldots l_k}$; to simplify notation I will write γ_ϕ and s_ϕ for such a property as $\gamma^{i_1 \ldots i_k}_{l_1 \ldots l_k}$ and $s^{i_1 \ldots i_k}_{l_1 \ldots l_k}$ respectively.[1]

If $A = \{i_1, \ldots, i_k\}$ then by $\mathcal{F}^A$ I will mean $\mathcal{F}^{i_1 \ldots i_k}$. Also, if $A_1, \ldots, A_j$ are disjoint subsets of $\{1, \ldots, n\}$ then $\mathcal{F}^{A_1 \ldots A_j}$ will be used as an abbreviation for $\mathcal{F}^{A_1 \cup \ldots \cup A_j}$. The notation F^A_L will denote an arbitrary property in $\mathcal{F}^A$ (so L is here a k-tuple of elements of $\{1, 2\}$); the notations γ^A_L and s^A_L are to be understood similarly. This notation is used in the following theorem. Proofs of all theorems are given in Section 8.

THEOREM 1. Let A and B be non-empty disjoint subsets of $\{1, \ldots, n\}$. If $\mathcal{F}^{AB}$ is a λ-family with respect to p then $\mathcal{F}^A$ is also a λ-family with respect to p, λ is the same for both families, and

$$\gamma^A_L = \sum_M \gamma^{AB}_{LM}.$$

As a simple example of its application, this theorem implies that if $\mathcal{F}^{12}$ is a λ-family then so is $\mathcal{F}^1$ and $\gamma^1_l = \gamma^{12}_{l1} + \gamma^{12}_{l2}$. Since $\mathcal{F}^{AB} = \mathcal{F}^{BA}$ the theorem likewise implies that $\mathcal{F}^2$ is a λ-family and $\gamma^2_l = \gamma^{12}_{1l} + \gamma^{12}_{2l}$.

3. $\mathcal{F}^{1 \ldots n}$ AS A λ-FAMILY

Carnap (1952) proposed a probability model which, applied to $\mathcal{X}$, consists of the probability functions on $\mathcal{X}$ in which $\mathcal{F}^{1 \ldots n}$ is a λ-family with each $\gamma^{1 \ldots n}_{l_1 \ldots l_n} = 1/2^n$.

By Theorem 1, these conditions imply that each $\mathcal{F}^i$ is a λ-family with $\gamma^i_1 = \gamma^i_2 = 1/2$. Hence by Definition 1, $p(F^i_1 a) = p(F^i_2 a) = 1/2$. However, if F^1_1 is the property Swan then, since this is just one of many things that an individual might be, $p(F^1_1 a)$ should be less than 1/2. Similarly, if F^2_1 is Australian and F^3_1 is White, $p(F^2_1 a)$ and $p(F^3_1 a)$ should be less than 1/2. This objection can be met by simply dropping the condition that $\gamma^{1 \ldots n}_{l_1 \ldots l_n} = 1/2^n$ and thus requiring only that $\mathcal{F}^{1 \ldots n}$ be a λ-family. I will use P_λ to denote the class of probability functions on $\mathcal{X}$ that satisfy this condition.

A probability function p that is properly sensitive to analogy will satisfy the following condition:[2]

$$p(F_1^3 a | F_{11}^{12} a \cap F_{121}^{123} b) > p(F_1^3 a | F_{11}^{12} a \cap F_{122}^{123} b).$$

Applied to my swan example, this says that the probability of an Australian swan being white is greater if a white non-Australian swan is observed than if the non-Australian swan had been non-white. However, for every $p \in P_\lambda$ we have

$$\begin{aligned} p(F_1^3 a | F_{11}^{12} a \cap F_{121}^{123} b) &= \frac{p(F_{111}^{123} a | F_{121}^{123} b)}{p(F_{111}^{123} a | F_{121}^{123} b) + p(F_{112}^{123} a | F_{121}^{123} b)} \\ &= \frac{\gamma_{111}^{123}}{\gamma_{111}^{123} + \gamma_{112}^{123}} \qquad \text{by Definition 1} \\ &= \frac{p(F_{111}^{123} a | F_{122}^{123} b)}{p(F_{111}^{123} a | F_{122}^{123} b) + p(F_{112}^{123} a | F_{122}^{123} b)}, \\ &\qquad \text{by Definition 1} \\ &= p(F_1^3 a | F_{11}^{12} a \cap F_{122}^{123} b). \end{aligned}$$

More generally, it can be shown that probability functions in P_λ are insensitive to all analogies between individuals whenever the individuals are known to differ in any way. Thus no $p \in P_\lambda$ is properly sensitive to analogy and so $R \cap P_\lambda = \emptyset$. For the special case in which $\gamma_{l_1 \ldots l_n}^{1 \ldots n} = 1/2^n$, this problem was discovered by Carnap in the early 1950s, apparently even before his (1952) appeared in print (Carnap, 1963, 974n.); later Achinstein (1963, 215ff.) independently made the same point.

4. THE MODEL FOR TWO PROPERTIES

After discovering the problem with analogy just mentioned, Carnap sought a new probability model that would be sensitive to analogies between individuals that are known to differ. Carnap initially developed a model for the case in which there are two basic families of properties (Carnap and Stegmüller 1959, 251ff.; Carnap 1975, 318ff.). Carnap allowed the basic families to contain any finite number of properties, but I am here considering only the case in which the basic families contain two properties. (The reason for this restriction is to avoid the need to consider analogy effects due to similarity relations between the properties within a basic family.)

Before describing Carnap's model it will be useful to have the following definition:

DEFINITION 2. Families of properties $\Phi^1, \ldots, \Phi^k$ are *probabilistically independent* in p iff the following holds: If $E^1, \ldots, E^k$ are sample propositions with respect to $\Phi^1, \ldots, \Phi^k$ respectively, and if each of $E^1, \ldots, E^k$ involves the same individuals, then

$$p(E^1 \cap \ldots \cap E^k) = p(E^1) \ldots p(E^k).$$

As a simple illustration, if $\mathcal{F}^1$ and $\mathcal{F}^{23}$ are probabilistically independent in p then $p(F^{123}_{l_1 l_2 l_3} a) = p(F^1_{l_1} a) p(F^{23}_{l_2 l_3} a)$.

Carnap's model for the case $n = 2$ can now be described as follows: Each probability function in the model is a mixture of two probability functions. One of these, which I will denote p^{12}, is a probability function in which $\mathcal{F}^{12}$ is a λ-family. The other, which I will denote $p^{1|2}$, is a probability function in which $\mathcal{F}^1$ and $\mathcal{F}^2$ are probabilistically independent λ-families with the same λ and γ values as in p^{12}. Carnap also required that $\gamma^{12}_{lm} = 1/4$ (in p^{12}), whence $\gamma^1_l = \gamma^2_m = 1/2$ (in p^{12} and $p^{1|2}$). He denoted the weight on $p^{1|2}$ by η, $0 < \eta < 1$. Thus for each p in Carnap's model we have

$$p = \eta p^{1|2} + (1 - \eta) p^{12}.$$

Maher (2000) proposed necessary conditions on R for the $n = 2$ case, identified the probability model defined by these conditions, and showed that this model is a generalization of Carnap's model for the same case. In this generalized model, Carnap's requirement that $\gamma^{12}_{lm} = 1/4$ is replaced by the weaker requirement that $\gamma^{12}_{lm} = \gamma^1_l \gamma^2_m$.

Maher (2000) proceeds as follows: The proposition that $\mathcal{F}^1$ and $\mathcal{F}^2$ are statistically independent (independent in physical probabilities or chances) is denoted I. De Finetti's representation theorem is used to define $p(I)$ and more generally $p(E \cap I)$ for any sample proposition E with respect to $\mathcal{F}^{12}$. Axioms governing probabilities conditional on I and $\bar{I}$ (the negation or complement of I) are stated; these axioms imply that $p(\cdot|\bar{I})$ is a probability function in which $\mathcal{F}^{12}$ is a λ-family, while $p(\cdot|I)$ is a probability function in which $\mathcal{F}^1$ and $\mathcal{F}^2$ are probabilistically independent λ-families with λ and γ^i_l the same as in p^{12}. By the law of total probability,

$$p(\cdot) = p(\cdot|I) p(I) + p(\cdot|\bar{I}) p(\bar{I}).$$

Thus Carnap's $p^{1|2}$ can be interpreted as $p(\cdot|I)$, his p^{12} as $p(\cdot|\bar{I})$, and his η as $p(I)$. Following Maher (2000), I will refer to this model as P_I.

Maher (2000) defended the adequacy of P_I and argued that it compared favorably with a variety of other models that could be applied to the problem of two properties. It is therefore of interest to consider how P_I might be generalized to cover cases in which $n > 2$; that question will be my focus in the remainder of this paper.

5. HESSE'S n-PROPERTY MODEL

Hesse (1964, 325) observed that the simplest way to generalize Carnap's two-property model to n properties is to set:

(1) $$p = \eta p^{1|\ldots|n} + (1-\eta)p^{1\ldots n}.$$

Here $p^{1\ldots n}$ is a probability function in which $\mathcal{F}^{1\ldots n}$ is a λ-family, while $p^{1|\ldots|n}$ is a probability function in which $\mathcal{F}^1, \ldots, \mathcal{F}^n$ are probabilistically independent λ-families with λ and γ^i_l being the same as for $\mathcal{F}^{1\ldots n}$. Hesse also followed Carnap in requiring that $\gamma^{1\ldots n}_{l_1\ldots l_n} = 1/2^n$ (in $p^{1\ldots n}$).

Hesse showed that all p in her model have the following desirable properties with regard to analogy:

$$\begin{aligned} p(F^3_1 a|F^{12}_{11}a \cap F^{123}_{111}b) &> p(F^3_1 a|F^{12}_{11}a \cap F^{123}_{211}b) \\ &> p(F^3_1 a|F^{12}_{11}a \cap F^{123}_{221}b) \\ &> p(F^3_1 a|F^{12}_{11}a \cap F^{123}_{222}b). \end{aligned}$$

The first of these inequalities also holds for all $p \in P_\lambda$; the other two do not because no $p \in P_\lambda$ takes account of analogies between individuals that are known to differ in any property.

Since Hesse's requirement that $\gamma^{1\ldots n}_{l_1\ldots l_n} = 1/2^n$ is unduly restrictive, in what follows I will consider a generalized version of her model in which this requirement is replaced by the weaker condition that $\gamma^{1\ldots n}_{l_1\ldots l_n} = \gamma^1_{l_1}\ldots\gamma^n_{l_n}$. I will refer to this generalized model as P_H.

5.1. *Foundation*

The foundation provided by Maher (2000) for P_I can be generalized to give a foundation for P_H, as I will now show.

Let Q denote the set of probability functions q on $\mathcal{X}$ such that, for any sample proposition E with respect to $\mathcal{F}^{1\ldots n}$ and any individual a, if $s^{1\ldots n}_{l_1\ldots l_n}$ is the number of individuals to which E ascribes $F^{1\ldots n}_{l_1\ldots l_n}$, then

$$q(E) = \prod_{l_i \in \{1,2\}} q(F^{1\ldots n}_{l_1\ldots l_n}a)^{s^{1\ldots n}_{l_1\ldots l_n}}.$$

Thus if $q \in Q$ then q has the properties one expects of physical probabilities or chances; in particular, $q(F^{1\ldots n}_{l_1\ldots l_n}a|E)$ is the same for all individuals a

not involved in the sample proposition E and is independent of E. We can think of any $A \subset Q$ as representing the proposition that the true chance distribution is in A; that thought motivates the following definition.

DEFINITION 3. The proposition that families of properties $\Phi^1, \dots, \Phi^k$ are *statistically independent*, denoted $\mathrm{Ind}(\Phi^1, \dots, \Phi^k)$, is the set of all $q \in Q$ which are such that, for any $\phi_i \in \Phi^i$, $i = 1, \dots, k$, and for any individual a,

$$q(\phi_1 a \cap \dots \cap \phi_k a) = q(\phi_1 a) \dots q(\phi_k a).$$

It is convenient to allow this definition to apply even in the degenerate case where $k = 1$, so that $\mathrm{Ind}(\Phi)$ holds trivially for any single family Φ.

The next definition generalizes the propositions I and $\bar{I}$ of Section 4. In this definition, the overbar again denotes negation or complementation.

DEFINITION 4. A *partition* of set S is a class of non-empty pairwise disjoint sets whose union is S. If $\mathcal{A} = \{A_1, \dots, A_k\}$ is a partition of $\{1, \dots, n\}$ then $I^{\mathcal{A}}$ denotes that

1. $\mathrm{Ind}(\mathcal{F}^{A_1}, \dots, \mathcal{F}^{A_k})$ and
2. for all $i = 1, \dots, k$ and all partitions $\{B_1, \dots, B_m\}$ of A_i, $m > 1$, $\overline{\mathrm{Ind}}(\mathcal{F}^{B_1}, \dots, \mathcal{F}^{B_m})$.

$I^{\mathcal{A}}$ will be called an *I-proposition*.

For brevity I will, when writing particular I-propositions, represent a partition by writing the members of each element of the partition separated by a vertical bar. For example, $I^{1|23}$ is short for $I^{\{\{1\},\{2,3\}\}}$, which means that $\mathrm{Ind}(\mathcal{F}^1, \mathcal{F}^{23})$ and $\overline{\mathrm{Ind}}(\mathcal{F}^2, \mathcal{F}^3)$. For another example, I^{123} is short for $I^{\{\{1,2,3\}\}}$ and means that $\overline{\mathrm{Ind}}(\mathcal{F}^1, \mathcal{F}^2, \mathcal{F}^3)$ and, for all distinct $i, j, k \in \{1, 2, 3\}$, $\overline{\mathrm{Ind}}(\mathcal{F}^i, \mathcal{F}^{jk})$. If $n = 2$ then $I^{1|2}$ and I^{12} are the same as the propositions I and $\bar{I}$, respectively, of Section 4.

THEOREM 2. Exactly one I-proposition is true.

P_H can now be derived in the following way: Assume that $I^{1\dots n}$ and $I^{1|\dots|n}$ are the only I-propositions with positive probability. Generalized versions of the axioms of Maher (2000) – which I will not state here – then imply that $p(\cdot|I^{1\dots n})$ is a probability function in which $\mathcal{F}^{1\dots n}$ is a λ-family, while $p(\cdot|I^{1|\dots|n})$ is a probability function in which $\mathcal{F}^1, \dots, \mathcal{F}^n$ are probabilistically independent λ-families with λ and γ_l^i the same as in $p(\cdot|I^{1\dots n})$. Then Hesse's $p^{1\dots n}$ would be interpreted as $p(\cdot|I^{1\dots n})$, her $p^{1|\dots|n}$

would be interpreted as $p(\cdot|I^{1|\dots|n})$, and her η would be identified with $p(I^{1|\dots|n})$.

However, this axiomatic derivation is implausible. An obvious objection is this: For any $n > 2$ there are other I-propositions besides $I^{1\dots n}$ and $I^{1|\dots|n}$; for example, if $n = 3$ the I-propositions are I^{123}, $I^{1|23}$, $I^{2|13}$, $I^{3|12}$, and $I^{1|2|3}$. To obtain P_H we had to suppose that only the first and last of these have positive probability, but the others are equally worthy of some credence.

5.2. *Predictive Properties*

The preceding discussion suggests that P_H corresponds to an inadequate view of the possible statistical independence relations. If that is right then the inadequacy should manifest itself in applications. I will now describe one way in which this happens.

Suppose that we know, of each individual in some sample, whether it is Australian and whether it is a swan but not whether it is white. The following theorem shows that, given any evidence of this kind, the further evidence that a non-Australian swan is white still confirms that an Australian swan is white.

THEOREM 3. Let $p \in P_H$ with $n > 2$ and let E^{12} be a sample proposition with respect to $\mathcal{F}^{12}$ that does not involve a or b. Then

$$p(F_1^3 a|F_{11}^{12} a \cap F_{121}^{123} b \cap E^{12}) > p(F_1^3 a|F_{11}^{12} a \cap F_{122}^{123} b \cap E^{12}).$$

This is as it should be. However, the amount of the confirmation approaches zero as the sample size approaches infinity if, in the samples, the following all eventually become and remain greater than some positive value: (1) The proportion of individuals that are Australian; (2) the proportion of individuals that are not Australian; and (3) the difference between the proportion of Australian individuals that are swans and the proportion of non-Australian individuals that are swans. This is stated more formally by the following theorem.

THEOREM 4. Let $\{E_{(s)}^{12}\}_{s=1}^{\infty}$ be a sequence of sample propositions with respect to $\mathcal{F}^{12}$, where each $E_{(s)}^{12}$ is for a sample of size s and $E_{(s+1)}^{12}$ entails $E_{(s)}^{12}$. Let s_L^A denote the number of individuals that are ascribed F_L^A by $E_{(s)}^{12}$. If there exists $\varepsilon > 0$ and integer S such that for all $s > S$

$$\frac{s_1^2}{s} > \varepsilon, \quad \frac{s_2^2}{s} > \varepsilon, \quad \text{and} \quad \left|\frac{s_{11}^{12}}{s_1^2} - \frac{s_{12}^{12}}{s_2^2}\right| > \varepsilon$$

then for any distinct a and b not involved in any $E^{12}_{(s)}$ and $p \in P_H$ with $n > 2$:

$$\lim_{s\to\infty} \left| p(F^3_1 a | F^{12}_{11} a \cap F^{123}_{121} b \cap E^{12}_{(s)}) - p(F^3_1 a | F^{12}_{11} a \cap F^{123}_{122} b \cap E^{12}_{(s)}) \right| = 0.$$

This seems quite wrong. The color of a non-Australian swan should not become practically irrelevant to the color of an Australian swan just because we know that, in a large sample, swans were less (or more) common in Australia than elsewhere.

If we assume the foundation for P_H in terms of I-propositions then the reason for this unsatisfactory result can be explained as follows: The information that the proportion of swans is different in large samples of Australian and non-Australian individuals makes it practically certain that $\overline{\text{Ind}}(\mathcal{F}^1, \mathcal{F}^2)$ and so $I^{1|\dots|n}$ is false. Since in Hesse's model $I^{1\dots n}$ is the only other possible I-proposition, p becomes practically equivalent to $p(\cdot|I^{1\dots n})$. But $p(\cdot|I^{1\dots n})$, which is Hesse's $p^{1\dots n}$, is a probability function in which $\mathcal{F}^{1\dots n}$ is a λ-family and thus it gives no analogy effect for individuals that are known to differ in any way. My proof of Theorem 4 follows essentially this reasoning but without assuming the foundation in terms of I-propositions.

6. CARNAP AND KEMENY'S n-PROPERTY MODEL

In the early 1950s Carnap and Kemeny jointly developed a generalization of Carnap's two-property model to n properties. This work has not been published but it is described in an unpublished document by Carnap (1954). For $n = 3$ their model is:

$$p = c^{123} p^{123} + c^{1|23} p^{1|23} + c^{12|3} p^{12|3} + c^{13|2} p^{13|2} + c^{1|2|3} p^{1|2|3}.$$

Here the $c^{\cdots}$ are positive constants that sum to one, generalizing the earlier η and $1 - \eta$. The probability functions p^{123} and $p^{1|2|3}$ are the same as for Hesse; in addition we have here probability functions of the form $p^{ij|k}$, in which $\mathcal{F}^{ij}$ and $\mathcal{F}^k$ are probabilistically independent λ-families with the same λ and γ values as in p^{123}.

If $\mathcal{A} = \{A_1, \dots, A_k\}$ is a partition of $\{1, \dots, n\}$, let $p^{\mathcal{A}}$ be a probability function in which $\mathcal{F}^{A_1}, \dots, \mathcal{F}^{A_k}$ are probabilistically independent λ-families with the same λ and γ values as in $p^{1\dots n}$. Then Carnap and Kemeny's model for any $n \geq 1$ is

$$p = \sum_{\mathcal{A}} c^{\mathcal{A}} p^{\mathcal{A}},$$

the summation being taken over all partitions of $\{1, \ldots, n\}$ and the $c^{\mathcal{A}}$ being positive constants that sum to one.

In giving this account I have not followed Carnap's 1954 notation. Also I have described the model for the case in which the basic families contain only two properties, though Carnap and Kemeny did not make this restriction. Carnap and Kemeny required that each $\gamma^{1\ldots n}_{l_1\ldots l_n} = 1/2^n$; as I did with earlier models, I will replace this with the weaker condition that $\gamma^{1\ldots n}_{l_1\ldots l_n} = \gamma^1_{l_1} \ldots \gamma^n_{l_n}$. I will refer to the resulting model as P_K.

So far as I know, Carnap and Kemeny did not investigate the analogical properties of their model.

6.1. *Foundation*

In Section 5.1 I indicated how P_H can be given a foundation using I-propositions by assuming that $I^{1\ldots n}$ and $I^{1|\cdots|n}$ are the only I-propositions with positive initial probability. If we replace that assumption with the more plausible condition that all I-propositions have positive initial probability, keeping everything else the same, we get a foundation for P_K. Thus Carnap and Kemeny's $p^{\mathcal{A}}$ can be interpreted as $p(\cdot|I^{\mathcal{A}})$ and their $c^{\mathcal{A}}$ can be interpreted as $p(I^{\mathcal{A}})$.

I will now argue that this foundation has several flaws. In exhibiting these flaws I will, for definiteness, consider only the case $n = 3$.

First flaw: Since $\mathcal{F}^{123}$ is a λ-family with respect to $p(\cdot|I^{123})$, it follows from Theorem 1 that $\mathcal{F}^{12}$ is also a λ-family with respect to $p(\cdot|I^{123})$. According to the foundation that I have sketched, this is appropriate if and only if it is given that $\mathcal{F}^1$ and $\mathcal{F}^2$ are statistically dependent. However, it is possible for I^{123} to be true and yet $\mathcal{F}^1$ and $\mathcal{F}^2$ to be statistically independent. What I^{123} asserts is $\overline{\text{Ind}}(\mathcal{F}^1, \mathcal{F}^2, \mathcal{F}^3)$ and $\overline{\text{Ind}}(\mathcal{F}^{ij}, \mathcal{F}^k)$ for all distinct $i, j, k \in \{1, 2, 3\}$, and this implies nothing about the truth values of the pairwise relations $\text{Ind}(\mathcal{F}^i, \mathcal{F}^j)$. For example, if

$$q^{123}_{111} = q^{123}_{122} = q^{123}_{212} = q^{123}_{221} = 1/16,$$
$$q^{123}_{112} = q^{123}_{121} = q^{123}_{211} = q^{123}_{222} = 3/16$$

then we have I^{123} but $\text{Ind}(\mathcal{F}^1, \mathcal{F}^2)$, $\text{Ind}(\mathcal{F}^1, \mathcal{F}^3)$, and $\text{Ind}(\mathcal{F}^2, \mathcal{F}^3)$.

Second flaw: If $\overline{\text{Ind}}(\mathcal{F}^1, \mathcal{F}^2)$ and $\overline{\text{Ind}}(\mathcal{F}^1, \mathcal{F}^3)$ then $I^{1|2|3}$, $I^{1|23}$, $I^{2|13}$, and $I^{3|12}$ are not possible and so I^{123} must hold. The reasoning used in the preceding paragraph shows that $\mathcal{F}^2$ and $\mathcal{F}^3$ are treated as statistically dependent given I^{123}. However, $\overline{\text{Ind}}(\mathcal{F}^1, \mathcal{F}^2)$ and $\overline{\text{Ind}}(\mathcal{F}^1, \mathcal{F}^3)$ do not entail $\overline{\text{Ind}}(\mathcal{F}^2, \mathcal{F}^3)$. For example, if

$$q^{123}_{121} = q^{123}_{112} = q^{123}_{122} = q^{123}_{211} = 1/16,$$
$$q^{123}_{111} = q^{123}_{221} = q^{123}_{212} = q^{123}_{222} = 3/16$$

then we have $\overline{\mathrm{Ind}}(\mathcal{F}^1, \mathcal{F}^2)$ and $\overline{\mathrm{Ind}}(\mathcal{F}^1, \mathcal{F}^3)$ but $\mathrm{Ind}(\mathcal{F}^2, \mathcal{F}^3)$.

Third flaw: If $\mathcal{F}^2$ and $\mathcal{F}^3$ are statistically dependent this does not settle whether they are statistically dependent given F^1_l. For example, if the chance of an Australian individual being white is less than that of a non-Australian individual, it does not follow that the chance of an Australian *swan* being white is less than that of a non-Australian swan. Mathematically, $\mathcal{F}^2$ and $\mathcal{F}^3$ are dependent iff $q^{23}_{l_2l_3} = q^2_{l_2}q^3_{l_3}$; they are dependent given $F^1_{l_1}$ iff

$$\frac{q^{123}_{l_1l_2l_3}}{q^1_{l_1}} = \frac{q^{12}_{l_1l_2}}{q^1_{l_1}}\frac{q^{13}_{l_1l_3}}{q^1_{l_1}};$$

and these are not equivalent conditions. Since the I-propositions merely represent overall statistical dependence or independence of families they do not distinguish the different possibilities for conditional dependence or independence. For example, it is not difficult to show that for any sample data E,

$$p(F^{23}_{l_2l_3}a|F^1_{l_1}a \cap E \cap I^{1|23}) = p(F^{23}_{l_2l_3}a|E \cap I^{1|23}).$$

Thus, given $I^{1|23}$, $\mathcal{F}^{23}$ is treated as a λ-family even given that the individual is $F^1_{l_1}$; this is appropriate only if $\mathcal{F}^2$ and $\mathcal{F}^3$ are dependent given $F^1_{l_1}$.

6.2. *Predictive Properties*

Since P_K corresponds to a more adequate view of the possible statistical relevance relations than P_H does, it correctly handles some applications that are mishandled by P_H. In particular, Theorem 4 does not hold for $p \in P_K$ and so the criticism of P_H that I made in Section 5.2 does not apply to P_K. Nevertheless, we have seen that the foundation for P_K in terms of I-propositions still fails to allow for some relations that are in fact possible. I will now describe one way in which this inadequacy can show up in applications.

Suppose we know, of each individual in some sample, whether it is Australian and whether it is white but not whether it is a swan. Suppose further that, as the sample size increases, the following all eventually become and remain larger than some positive value: (1) The proportion of individuals that are Australian; (2) the proportion of individuals that are not Australian; (3) the difference between the proportion of Australian individuals that are white and the proportion of non-Australian individuals that are white. Then, in the limit as the sample size approaches infinity, the further evidence that a non-Australian swan is white does not confirm that an Australian swan is white. This is stated more formally by the following theorem:

THEOREM 5. Let $\{E^{23}_{(s)}\}^{\infty}_{s=1}$ be a sequence of sample propositions with respect to $\mathcal{F}^{23}$, where each $E^{23}_{(s)}$ is for a sample of size s and $E^{23}_{(s+1)}$ entails $E^{23}_{(s)}$. Let s^A_L denote the number of individuals that are ascribed F^A_L by $E^{23}_{(s)}$. If there exists $\varepsilon > 0$ and integer S such that for all $s > S$

$$\frac{s^2_1}{s} > \varepsilon, \quad \frac{s^2_2}{s} > \varepsilon, \quad \text{and} \quad \left|\frac{s^{23}_{11}}{s^2_1} - \frac{s^{23}_{21}}{s^2_2}\right| > \varepsilon$$

then for any distinct a and b not involved in any $E^{23}_{(s)}$ and $p \in P_K$ with $n > 2$:

$$\lim_{s\to\infty} \left|p(F^3_1 a|F^{12}_{11}a \cap F^{123}_{121}b \cap E^{23}_{(s)}) - p(F^3_1 a|F^{12}_{11}a \cap F^{123}_{122}b \cap E^{23}_{(s)})\right| = 0.$$

This seems quite wrong. The evidence here indicates that the proportion of white things in Australia is different to elsewhere, but it does not imply that Australian swans differ from non-Australian swans in color; hence this evidence is not a reason to deem the color of non-Australian swans irrelevant to the color of Australian swans.

This unsatisfactory result is due to the last of the three flaws that I noted in the foundation for P_K using I-propositions. In terms of this foundation, what the evidence $E^{23}_{(s)}$ does is make it practically certain that $\mathcal{F}^2$ and $\mathcal{F}^3$ are statistically dependent. Then $\mathcal{F}^2$ and $\mathcal{F}^3$ are also treated as dependent given F^1_1 and so, even for individuals known to be F^1_1, there is no analogy effect when the individuals are known to differ in some way.

The first two flaws that I noted in the foundation for P_K can also show up in applications, but one counterexample is enough.

7. CONCLUSION

Following my (2000) defense of the probability model P_I, which is a generalization of Carnap's model for two properties, it is natural to ask how P_I could be extended to deal with more than two properties. To my knowledge, only two such generalizations have been proposed: Hesse's simple model P_H and Carnap and Kemeny's more elaborate P_K. But I have argued that, when $n > 2$, every $p \in P_H$ and $p \in P_K$ fails to properly reflect correct analogical reasoning; hence $R \cap P_H = \emptyset$ and $R \cap P_K = \emptyset$.

The foundation in terms of I-propositions suggests that in each case the underlying reason for the failures is that both models give zero probability to some possible patterns of statistical dependence relations between the basic families of properties. It thus appears that a satisfactory generalization of P_I must be more complex than even P_K; what form such a model should take is question for future research.

8. PROOFS

8.1. *Proof of Theorem 1*

Let $\mathcal{F}^{AB}$ be a λ-family with respect to p. Let E^A be a sample proposition with respect to $\mathcal{F}^A$ and let E^{AB} be a sample proposition with respect to $\mathcal{F}^{AB}$ that involves the same individuals as E^A and is such that $E^{AB} \subset E^A$. Thus s^A_L, the number of individuals having F^A_L, is the same in E^{AB} and E^A. Then for any individual a not involved in E:

$$\begin{aligned} p(F^A_L a | E^{AB}) &= \sum_M p(F^{AB}_{LM} a | E^{AB}) \\ &= \sum_M \frac{s^{AB}_{LM} + \lambda\gamma^{AB}_{LM}}{s+\lambda} \qquad \text{by Definition 1} \\ &= \frac{s^A_L + \lambda \sum_M \gamma^{AB}_{LM}}{s+\lambda}. \end{aligned}$$

Since this holds for every E^{AB} and the union of all of them is E^A, it follows from the law of total probability that

$$p(F^A_L a | E^A) = \frac{s^A_L + \lambda \sum_M \gamma^{AB}_{LM}}{s+\lambda}.$$

This is what the theorem asserts.

8.2. *Proof of Theorem 2*

If $n = 1$ then $\{\{1\}\}$ is the only partition of $\{1, \ldots, n\}$ and so I^1 is the only I-proposition. Also I^1 is trivially true, so the theorem holds. In what follows I assume that $n > 1$.

I will first prove that the I-propositions are exhaustive. Suppose that $\bar{I}^{\mathcal{A}}$ holds for all partitions $\mathcal{A}$ of $\{1, \ldots, n\}$ other than the trivial partition $\{\{1, \ldots, n\}\}$. I will show that in this case $I^{1\ldots n}$ holds.

Let $S(k)$ denote that $\overline{\text{Ind}}(\mathcal{F}^{A_1}, \ldots, \mathcal{F}^{A_l})$ for all partitions $\{A_1, \ldots, A_l\}$ of $\{1, \ldots, n\}$, $l \geq k$. By assumption $\bar{I}^{1|2|\ldots|n}$ and so, by Definition 4,

$\overline{\text{Ind}}(\mathcal{F}^1, \ldots, \mathcal{F}^n)$. Thus $S(n)$ holds. Now suppose $S(k)$ holds for some $k \in \{3, \ldots, n\}$ and let $\mathcal{A} = \{A_1, \ldots, A_{k-1}\}$ be a partition of $\{1, \ldots, n\}$. If $\text{Ind}(\mathcal{F}^{A_1}, \ldots, \mathcal{F}^{A_{k-1}})$ then, since $\bar{I}^{A_1|\ldots|A_{k-1}}$, there exists $i \in \{1, \ldots, k-1\}$ and a partition $\{B_1, \ldots, B_m\}$ of A_i, $m > 1$, such that $\text{Ind}\{\mathcal{F}^{B_1}, \ldots, \mathcal{F}^{B_m}\}$. It follows that

$$\text{Ind}(\mathcal{F}^{A_1}, \ldots, F^{A_{i-1}}, F^{B_1}, \ldots, F^{B_m}, F^{A_{i+1}}, \ldots, F^{A_{k-1}}).$$

Since $m > 1$ this contradicts $S(k)$. Hence $\overline{\text{Ind}}(\mathcal{F}^{A_1}, \ldots, \mathcal{F}^{A_{k-1}})$. Thus $S(k-1)$ is true. So by mathematical induction, $S(k)$ is true for all $k = 2, \ldots, n$. Since $\text{Ind}(\mathcal{F}^{1\ldots n})$ is trivially true it follows that $I^{1\ldots n}$. Hence the I-propositions are exhaustive.

I will now prove that the I-propositions are pairwise exclusive. Let $\mathcal{A}$ and $\mathcal{B}$ be different partitions of $\{1, \ldots, n\}$ and suppose $I^{\mathcal{A}}$ and $I^{\mathcal{B}}$. Since $\mathcal{A} \neq \mathcal{B}$ there exist $A \in \mathcal{A}$ and $B \in \mathcal{B}$ such that at least one of the following holds:

$$(2) \qquad \begin{aligned} &\emptyset \neq A \cap B \neq A, \\ &\emptyset \neq A \cap B \neq B. \end{aligned}$$

By reversing the labeling of $\mathcal{A}$ and $\mathcal{B}$ if necessary we can ensure that (2) holds and I will assume that this has been done. Let

$$\begin{aligned} A \cap B &= \{i_1, \ldots, i_\alpha\}, \\ A \cap \bar{B} &= \{i_{\alpha+1}, \ldots, i_\beta\}, \\ \bar{A} \cap B &= \{i_{\beta+1}, \ldots, i_\gamma\}, \\ \bar{A} \cap \bar{B} &= \{i_{\gamma+1}, \ldots, i_n\}. \end{aligned}$$

Then for all $q \in I^{\mathcal{B}}$ and any individual a we have:

$$\begin{aligned} q(F^{i_1\ldots i_\beta}_{l_1\ldots l_\beta} a) &= \sum_{l_{\beta+1},\ldots,l_n \in \{1,2\}} q(F^{i_1\ldots i_n}_{l_1\ldots l_n} a) \\ &= \sum_{l_{\beta+1},\ldots,l_n \in \{1,2\}} q(F^{i_1\ldots i_\alpha i_{\beta+1}\ldots i_\gamma}_{l_1\ldots l_\alpha l_{\beta+1}\ldots l_\gamma} a) q(F^{i_{\alpha+1}\ldots i_\beta i_{\gamma+1}\ldots i_n}_{l_{\alpha+1}\ldots l_\beta l_{\gamma+1}\ldots l_n} a) \\ &= q(F^{i_1\ldots i_\alpha}_{l_1\ldots l_\alpha} a) q(F^{i_{\alpha+1}\ldots i_\beta}_{l_{\alpha+1}\ldots l_\beta} a). \end{aligned}$$

Thus we have $\text{Ind}(\mathcal{F}^{A\cap B}, \mathcal{F}^{A\cap\bar{B}})$, which contradicts the assumption that $I^{\mathcal{A}}$. Hence the supposition from which we began, namely that $I^{\mathcal{A}}$ and $I^{\mathcal{B}}$ for different partitions $\mathcal{A}$ and $\mathcal{B}$ of $\{1, \ldots, n\}$, is false; so the I-propositions are pairwise exclusive.

8.3. *Lemmas Used in the Proof of Theorem 3*

The lemmas in this section have been stated in a more general form than is needed for proving Theorem 3 because they will also be used in the proof of Theorem 5.

In what follows, $\mathcal{X}^A$, for any non-empty $A \subset \{1, \ldots, n\}$, denotes the subalgebra of $\mathcal{X}$ obtained by using only the properties in $\mathcal{F}^A$ rather than the more specific properties in $\mathcal{F}^{1\ldots n}$.

LEMMA 1. Let p_1 and p_2 be two probability functions on $\mathcal{X}$ and let $A \subset \{1, \ldots, n\}$, $A \neq \emptyset$. If $\mathcal{F}^A$ is a λ-family with the same λ and γ values relative to both p_1 and p_2 then p_1 and p_2 agree on $\mathcal{X}^A$.

Proof. Let E^A be a sample proposition with respect to $\mathcal{F}^A$ and let s be the number of individuals involved in E^A. If $s = 0$ then $p_1(E^A) = p_2(E^A) = 1$.

Suppose now that for some $s \geq 0$, $p_1(E^A) = p_2(E^A)$ for all E^A involving s individuals. Let a be an individual not involved in E^A. Then

$$p_1(F_L^A a | E^A) = \frac{s_L^A + \lambda\gamma_L^A}{s + \lambda} = p_2(F_L^A a | E^A).$$

Thus

$$\frac{p_1(F_L^A a \cap E^A)}{p_1(E^A)} = \frac{p_2(F_L^A a \cap E^A)}{p_2(E^A)}.$$

Since $p_1(E^A) = p_2(E^A)$ it follows that

$$p_1(F_L^A a \cap E^A) = p_2(F_L^A a \cap E^A).$$

Thus p_1 and p_2 agree on all sample propositions with respect to $\mathcal{F}^A$ that involve $s + 1$ individuals. Hence by mathematical induction, p_1 and p_2 agree on all sample propositions with respect to $\mathcal{F}^A$.

Since every proposition in $\mathcal{X}^A$ is a disjoint union of sample propositions with respect to $\mathcal{F}^A$, it follows that $p_1(D) = p_2(D)$ for every $D \in \mathcal{X}^A$.

DEFINITION 5. Let $p^{1\ldots n}$ be a probability function with respect to which $\mathcal{F}^{1\ldots n}$ is a λ-family and let A be a non-empty subset of $\{1, \ldots, n\}$. Then $p^A \stackrel{\text{df}}{=}$ the restriction of $p^{1\ldots n}$ to $\mathcal{X}^A$. Also, if $\mathcal{A} = \{A_1, \ldots, A_k\}$ is a partition of A, then $p^{\mathcal{A}} \stackrel{\text{df}}{=}$ the probability function on $\mathcal{X}^A$ in which $\mathcal{F}^{A_1}, \ldots, \mathcal{F}^{A_k}$ are probabilistically independent λ-families with the same λ and γ values as in $p^{1\ldots n}$ (or, equivalently, p^A).

LEMMA 2. Let $\mathcal{A} = \{A_1, \dots, A_k\}$ be a partition of $\{1, \dots, n\}$ and let $B \subset \{1, \dots, n\}$, $B \neq \emptyset$. Let $\mathcal{A}_B = \{A_i \cap B : i = 1, \dots, k\} \setminus \{\emptyset\}$. Then if E^B is a sample proposition with respect to $\mathcal{F}^B$, $p^{\mathcal{A}}(E^B) = p^{\mathcal{A}_B}(E^B)$.

Proof. If $B = \{1, \dots, n\}$ then $\mathcal{A}_B = \mathcal{A}$ and the lemma holds trivially. So suppose B is a proper subset of $\{1, \dots, n\}$ and let $\bar{B} = \{1, \dots, n\} \setminus B$. For any $C \subset \{1, \dots, n\}$ let E^C denote a sample proposition with respect to $\mathcal{F}^C$ or, if $C = \emptyset$, let E^C be the necessarily true proposition. For $i \in \{1, \dots, n\}$ I will write $E^{\{i\}}$ simply as E^i. Then

$$\begin{aligned}(3) \qquad p^{\mathcal{A}}(E^B) &= \sum_{\{E^i : i \notin B\}} p^{\mathcal{A}}(E^B \cap (\cap_{i \notin B} E^i)) \\ &= \sum_{\{E^i : i \notin B\}} \prod_{j=1}^{k} p^{\mathcal{A}}(E^{A_j \cap B} \cap (\cap_{i \in A_j \setminus B} E^i)), \\ &\qquad \text{by Definition 5} \\ &= \prod_{j=1}^{k} p^{\mathcal{A}}(E^{A_j \cap B}).\end{aligned}$$

Terms $p^{\mathcal{A}}(E^{A_j \cap B})$ for which $A_j \cap B = \emptyset$ can be deleted from this last product without altering its value.

By Definition 5, each $\mathcal{F}^{A_j}$, $j = 1, \dots, k$, is a λ-family relative to $p^{\mathcal{A}}$ and the values of λ and $\gamma^{A_j}_{L_j}$ are the same as for $p^{1\dots n}$. Hence by Theorem 1, each $\mathcal{F}^{A_j \cap B}$, $A_j \cap B \neq \emptyset$, is a λ-family relative to $p^{\mathcal{A}}$ and has the same λ and γ values as for $p^{1\dots n}$. The same is true, by definition, for $p^{\mathcal{A}_B}$. Hence by Lemma 1,

$$p^{\mathcal{A}}(E^{A_j \cap B}) = p^{\mathcal{A}_B}(E^{A_j \cap B}), \quad j = 1, \dots, k \text{ and } A_j \cap B \neq \emptyset.$$

Substituting in (3) then gives:

$$\begin{aligned}p^{\mathcal{A}}(E^B) &= \prod_{j=1}^{k} p^{\mathcal{A}_B}(E^{A_j \cap B}) \\ &= p^{\mathcal{A}_B}(E^B) \qquad \text{by Definition 5.}\end{aligned}$$

In the rest of this section I use the convention that if E is a sample proposition with respect to $\mathcal{F}^{1\dots n}$ and $A \subset \{1, \dots, n\}$, $A \neq \emptyset$, then E^A denotes the sample proposition with respect to $\mathcal{F}^A$ that involves the same individuals as E and is entailed by E.

LEMMA 3. If $\mathcal{F}^{1...n}$ is a λ-family relative to p, D and E are sample propositions relative to $\mathcal{F}^{1...n}$, no individual is involved in both D and E, a is involved in neither D nor E, and $A \subset \{1, \dots, n\}$, $A \neq \emptyset$, then

$$p(F_L^A a | D \cap E^A) = p(F_L^A a | D \cap E).$$

Proof. Let C be a sample proposition with respect to $\mathcal{F}^{1...n}$ such that $C^A = E^A$. Let t be the number of individuals involved in $D \cap C$ and let $t_{LM}^{A\bar{A}}$ be the number of them that are ascribed $F_{LM}^{A\bar{A}}$ by $D \cap C$.

$$\begin{aligned} p(F_L^A a | D \cap C) &= \sum_M p(F_{LM}^{A\bar{A}} a | D \cap C) \\ &= \sum_M \frac{t_{LM}^{A\bar{A}} + \lambda \gamma_{LM}^{A\bar{A}}}{t + \lambda} \\ &= \frac{t_L^A + \lambda \gamma_L^A}{t + \lambda}. \end{aligned}$$

Since this does not depend on $t_M^{\bar{A}}$ it is the same for all C satisfying the stated conditions on C and in particular is true for E. Hence

(4) $$p(F_L^A a | D \cap C) = p(F_L^A a | D \cap E).$$

So

$$\begin{aligned} p(F_L^A a | D \cap E^A) &= \sum_C p(F_L^A a | D \cap C)\, p(C | D \cap E^A) \\ &= p(F_L^A a | D \cap E) \sum_C p(C | D \cap E^A), \text{ by (4)} \\ &= p(F_L^A a | D \cap E). \end{aligned}$$

LEMMA 4. If $\mathcal{F}^{1...n}$ is a λ-family with respect to p, D and E are sample propositions with respect to $\mathcal{F}^{1...n}$ that do not involve any of the same individuals, and $A \subset \{1, \dots, n\}$, $A \neq \emptyset$, then

(5) $$p(E^A | D) = p(E^A | D^A).$$

Proof. If the number of individuals involved in E is 0 then E is the necessarily true proposition and so $p(E^A | D) = p(E^A | D^A) = 1$, satisfying (5).

Now suppose that, for some $s \geq 0$, (5) holds for all E involving s individuals. Let E involve s individuals and let a be any individual not involved in D or E. Then

$$\begin{aligned} p(E^A \cap F_L^A a|D) &= p(E^A|D)\, p(F_L^A a|D \cap E^A) \\ &= p(E^A|D^A)\, p(F_L^A a|D \cap E^A) \qquad \text{by assumption} \\ &= p(E^A|D^A)\, p(F_L^A a|D \cap E) \qquad \text{by Lemma 3.} \end{aligned}$$

Applying Lemma 3 again to $p(F_L^A a|D \cap E)$, but this time with $D \cap E$ as the E of Lemma 3 (so that the D of Lemma 3 is the necessarily true proposition) we obtain:

$$\begin{aligned} p(E^A \cap F_L^A a|D) &= p(E^A|D^A)\, p(F_L^A a|D^A \cap E^A) \\ &= p(E^A \cap F_L^A a|D^A). \end{aligned}$$

Hence (5) holds for any E involving $s+1$ individuals. So by mathematical induction, (5) holds for all E.

LEMMA 5. Let $A \subset \{1, \ldots, n\}$, $A \neq \emptyset$. If $\mathcal{F}^A$ and $\mathcal{F}^{\bar{A}}$ are Probabilistically independent in p and if D and E are sample propositions with respect to $\mathcal{F}^{1\ldots n}$ that do not involve any of the same individuals then

(6) $$p(E^A|D) = p(E^A|D^A).$$

Proof. Let C denote any sample proposition with respect to $\mathcal{F}^{1\ldots n}$ such that $C^A = E^A$. Then

$$\begin{aligned} p(E^A|D) &= \sum_C p(C|D) \\ &= \sum_C \frac{p(C \cap D)}{p(D)} \\ &= \sum_C \frac{p(C^A \cap D^A)\, p(C^{\bar{A}} \cap D^{\bar{A}})}{p(D^A)\, p(D^{\bar{A}})} \qquad \text{by Definition 2} \\ &= \sum_C \frac{p(E^A \cap D^A)\, p(C^{\bar{A}} \cap D^{\bar{A}})}{p(D^A)\, p(D^{\bar{A}})} \\ &= p(E^A|D^A) \sum_C p(C^{\bar{A}}|D^{\bar{A}}) \\ &= p(E^A|D^A). \end{aligned}$$

8.4. *Proof of Theorem 3*

I will prove the theorem for the special case in which $n = 3$. It follows by Lemma 2 that the theorem also holds for $n > 3$.

$$p(F_1^3 a | F_{11}^{12} a \cap F_{121}^{123} b \cap E^{12}) = 1 \Bigg/ \left[1 + \frac{p(F_{112}^{123} a \cap F_{121}^{123} b \cap E^{12})}{p(F_{111}^{123} a \cap F_{121}^{123} b \cap E^{12})} \right],$$

$$p(F_1^3 a | F_{11}^{12} a \cap F_{122}^{123} b \cap E^{12}) = 1 \Bigg/ \left[1 + \frac{p(F_{112}^{123} a \cap F_{122}^{123} b \cap E^{12})}{p(F_{111}^{123} a \cap F_{122}^{123} b \cap E^{12})} \right].$$

Hence the theorem is true iff

$$(7) \qquad \frac{p(F_{111}^{123} a \cap F_{121}^{123} b \cap E^{12})}{p(F_{112}^{123} a \cap F_{121}^{123} b \cap E^{12})} > \frac{p(F_{111}^{123} a \cap F_{122}^{123} b \cap E^{12})}{p(F_{112}^{123} a \cap F_{122}^{123} b \cap E^{12})}.$$

Let

$$\alpha = (1 - \eta) \frac{\lambda \gamma_{11}^{12} \gamma_{12}^{12}}{1 + \lambda} p^{123}(E^{12} | F_{11}^{12} a \cap F_{12}^{12} b).$$

Then

$$\begin{aligned}
(8) \qquad & (1 - \eta) p^{123}(F_{111}^{123} a \cap F_{121}^{123} b \cap E^{12}) \\
&= (1 - \eta) p^{123}(F_{111}^{123} a) p^{123}(F_{121}^{123} b | F_{111}^{123} a) \cdot \\
&\quad p^{123}(E^{12} | F_{111}^{123} a \cap F_{121}^{123} b) \\
&= (1 - \eta) \gamma_{111}^{123} \frac{\lambda \gamma_{121}^{123}}{1 + \lambda} p^{123}(E^{12} | F_{111}^{123} a \cap F_{121}^{123} b) \\
&= (1 - \eta) \gamma_{111}^{123} \frac{\lambda \gamma_{121}^{123}}{1 + \lambda} p^{123}(E^{12} | F_{11}^{12} a \cap F_{12}^{12} b) \quad \text{by Lemma 4} \\
&= \alpha (\gamma_1^3)^2 \qquad \text{since } \gamma_{l_1 l_2 l_3}^{123} = \gamma_{l_1}^1 \gamma_{l_2}^2 \gamma_{l_3}^3 \text{ for } p \in P_H.
\end{aligned}$$

Similarly,

$$\begin{aligned}
(9) \qquad & (1 - \eta) p^{123}(F_{112}^{123} a \cap F_{121}^{123} b \cap E^{12}) = \alpha \gamma_1^3 \gamma_2^3, \\
(10) \qquad & (1 - \eta) p^{123}(F_{111}^{123} a \cap F_{122}^{123} b \cap E^{12}) = \alpha \gamma_1^3 \gamma_2^3, \\
(11) \qquad & (1 - \eta) p^{123}(F_{112}^{123} a \cap F_{122}^{123} b \cap E^{12}) = \alpha (\gamma_2^3)^2.
\end{aligned}$$

Let

$$\beta = \eta \frac{\gamma_1^1 (1 + \lambda \gamma_1^1) \gamma_1^2 \lambda \gamma_2^2}{(1 + \lambda)^3} p^{1|2|3}(E^{12} | F_{11}^{12} a \cap F_{12}^{12} b).$$

Then

$$
\begin{aligned}
(12)\quad & \eta p^{1|2|3}(F^{123}_{111}a \cap F^{123}_{121}b \cap E^{12}) \\
&= \eta p^{1|2|3}(F^{123}_{111}a)p^{1|2|3}(F^{123}_{121}b|F^{123}_{111}a)p^{1|2|3}(E^{12}|F^{123}_{111}a \cap F^{123}_{121}b) \\
&= \eta\gamma^1_1\gamma^2_1\gamma^3_1 \frac{(1+\lambda\gamma^1_1)\lambda\gamma^2_2(1+\lambda\gamma^3_1)}{(1+\lambda)^3} p^{1|2|3}(E^{12}|F^{123}_{111}a \cap F^{123}_{121}b) \\
&= \eta\gamma^1_1\gamma^2_1\gamma^3_1 \frac{(1+\lambda\gamma^1_1)\lambda\gamma^2_2(1+\lambda\gamma^3_1)}{(1+\lambda)^3} p^{1|2|3}(E^{12}|F^{12}_{11}a \cap F^{12}_{12}b), \\
&\qquad \text{by Lemma 5} \\
&= \beta\gamma^3_1(1+\lambda\gamma^3_1).
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
(13)\quad & \eta p^{1|2|3}(F^{123}_{112}a \cap F^{123}_{121}b \cap E^{12}) = \beta\gamma^3_1\lambda\gamma^3_2, \\
(14)\quad & \eta p^{1|2|3}(F^{123}_{111}a \cap F^{123}_{122}b \cap E^{12}) = \beta\gamma^3_1\lambda\gamma^3_2, \\
(15)\quad & \eta p^{1|2|3}(F^{123}_{112}a \cap F^{123}_{122}b \cap E^{12}) = \beta\gamma^3_2(1+\lambda\gamma^3_2).
\end{aligned}
$$

Hence

$$
\begin{aligned}
\frac{p(F^{123}_{111}a \cap F^{123}_{121}b \cap E^{12})}{p(F^{123}_{112}a \cap F^{123}_{121}b \cap E^{12})} &= \frac{\alpha(\gamma^3_1)^2 + \beta\gamma^3_1(1+\lambda\gamma^3_1)}{\alpha\gamma^3_1\gamma^3_2 + \beta\gamma^3_1\lambda\gamma^3_2} \\
&\qquad \text{by (8), (9), (12), and (13)} \\
&> \frac{\alpha(\gamma^3_1)^2 + \beta\lambda(\gamma^3_1)^2}{\alpha\gamma^3_1\gamma^3_2 + \beta\gamma^3_1\lambda\gamma^3_2} \\
&= \frac{\alpha\gamma^3_1\gamma^3_2 + \beta\gamma^3_1\lambda\gamma^3_2}{\alpha(\gamma^3_2)^2 + \beta\lambda(\gamma^3_2)^2} \\
&> \frac{\alpha\gamma^3_1\gamma^3_2 + \beta\gamma^3_1\lambda\gamma^3_2}{\alpha(\gamma^3_2)^2 + \beta\gamma^3_2(1+\lambda\gamma^3_2)} \\
&= \frac{p(F^{123}_{111}a \cap F^{123}_{122}b \cap E^{12})}{p(F^{123}_{112}a \cap F^{123}_{122}b \cap E^{12})}, \\
&\qquad \text{by (10), (11), (14), and (15).}
\end{aligned}
$$

Hence (7) holds and the theorem is proved.

8.5. *Lemmas Used in the Proof of Theorem 4*

It is convenient to be able to use the notation $p(I^{\mathcal{A}}|A)$ to denote the weight on $p^{\mathcal{A}}$ given A. To be able to do this without using any assumptions not made by Hesse or Carnap and Kemeny, I adopt the following definition.

DEFINITION 6. Let $p = \sum_{\mathcal{A}} c^{\mathcal{A}} p^{\mathcal{A}}$, where the summation is taken over all partitions $\mathcal{A}$ of $\{1, \ldots, n\}$ and the $c^{\mathcal{A}}$ are non-negative constants that sum to 1 (some of them may be zero). Then for any $A \in \mathcal{X}$ for which $p(A) > 0$, and any partition $\mathcal{A}$ of $\{1, \ldots, n\}$,

$$p(I^{\mathcal{A}}|A) \stackrel{\text{df}}{=} c^{\mathcal{A}} p^{\mathcal{A}}(A)/p(A).$$

LEMMA 6. Let $p = \sum_{\mathcal{A}} c^{\mathcal{A}} p^{\mathcal{A}}$, where the summation is taken over all partitions $\mathcal{A}$ of $\{1, \ldots, n\}$ and the $c^{\mathcal{A}}$ are non-negative constants that sum to 1 (some of them may be zero). Then if $A, B \in \mathcal{X}$ and $p(A) > 0$,

$$p(B|A) = \sum_{\mathcal{A}} p^{\mathcal{A}}(B|A)\, p(I^{\mathcal{A}}|A).$$

Proof.

$$\begin{aligned} p(B|A) &= \frac{p(A \cap B)}{p(A)} \\ &= \sum_{\mathcal{A}} \frac{c^{\mathcal{A}} p^{\mathcal{A}}(A \cap B)}{p(A)} \\ &= \sum_{\mathcal{A}} \frac{p^{\mathcal{A}}(A \cap B)\, p(I^{\mathcal{A}}|A)}{p^{\mathcal{A}}(A)} && \text{by Definition 6} \\ &= \sum_{\mathcal{A}} p^{\mathcal{A}}(B|A)\, p(I^{\mathcal{A}}|A). \end{aligned}$$

LEMMA 7. If $p \in P_H$ and $A, B \in \mathcal{X}$, with $p(A) > 0$, then

$$p(B|A) = p^{1|\ldots|n}(B|A)p(I^{1|\ldots|n}|A) + p^{1\ldots n}(B|A)p(I^{1\ldots n}|A).$$

Proof. We have $p = \sum_{\mathcal{A}} c^{\mathcal{A}} p^{\mathcal{A}}$ with $c^{\mathcal{A}} = 0$ for all $\mathcal{A}$ other than $\{\{1, \ldots, n\}\}$ and $\{\{1\}, \ldots, \{n\}\}$. Hence by Definition 6, $p(I^{\mathcal{A}}|A) = 0$ for $\mathcal{A}$ other than these two. This together with Lemma 6 gives the result.

The following definition generalizes Definition 1 by making it relative to a set of individuals V.

DEFINITION 7. A family of properties Φ is a λ*-family relative to probability function p for the individuals in V* $\stackrel{\text{df}}{=}$ there exists $\lambda > 0$ and for each $\phi \in \Phi$ there exists $\gamma_\phi \in (0, 1)$ such that the following holds: If E is a sample proposition with respect to Φ involving s individuals, all of which are in V, and if s_ϕ is the number of individuals to which E ascribes property ϕ, then for any individual $a \in V$ not involved in E,

$$p(\phi a|E) = \frac{s_\phi + \lambda\gamma_\phi}{s + \lambda}.$$

I will use the term *sample data* to refer to a proposition of the form

$$F_{L_1}^{A_1} a_1 \cap \ldots \cap F_{L_s}^{A_s} a_s.$$

Here the A_i are (not necessarily different) subsets of $\{1, \ldots, n\}$.

LEMMA 8. If $\mathcal{F}^{1\ldots n}$ is a λ-family relative to p and D is sample data then $\mathcal{F}^{1\ldots n}$ is a λ-family relative to $p(\cdot|D)$ for the individuals not involved in D.

Proof. Let t be the number of individuals involved in D. If $t = 0$ then the lemma is trivially true. Now suppose that the lemma holds for $t = \tau \geq 0$ and let $\tilde{p}(\cdot) = p(\cdot|D)$. By assumption, $\mathcal{F}^{1\ldots n}$ is a λ-family with respect to $\tilde{p}$ for individuals not involved in D; I will use $\tilde{\lambda}$ and $\tilde{\gamma}$ to denote parameters for $\tilde{p}$. Let E be a sample proposition with respect to $\mathcal{F}^{1\ldots n}$ that does not involve any individuals involved in D; I will use s to denote the number of individuals involved in E. Let a and b be distinct individuals not involved in D or E. Let $A \subset \{1, \ldots, n\}$, $F_L^A \in \mathcal{F}^A$, and $\xi_M = \tilde{p}(F_M^{\bar{A}} a|E \cap F_L^A a)$. Then

$$\begin{aligned}
\tilde{p}(F_{LM}^{A\bar{A}} b|E \cap F_L^A a) &= \sum_{M'} \tilde{p}(F_{LM}^{A\bar{A}} b|E \cap F_{LM'}^{A\bar{A}} a)\xi_{M'} \\
&= \xi_M \frac{s_{LM}^{A\bar{A}} + 1 + \tilde{\lambda}\tilde{\gamma}_{LM}^{A\bar{A}}}{s + 1 + \tilde{\lambda}} + \sum_{M' \neq M} \xi_{M'} \frac{s_{LM}^{A\bar{A}} + \tilde{\lambda}\tilde{\gamma}_{LM}^{A\bar{A}}}{s + 1 + \tilde{\lambda}} \\
&= \frac{s_{LM}^{A\bar{A}} + \xi_M + \tilde{\lambda}\tilde{\gamma}_{LM}^{A\bar{A}}}{s + 1 + \tilde{\lambda}} \qquad \text{since } \textstyle\sum_{M'} \xi_{M'} = 1 \\
&= \frac{s_{LM}^{A\bar{A}} + (1 + \tilde{\lambda})\frac{\xi_M + \tilde{\lambda}\tilde{\gamma}_{LM}^{A\bar{A}}}{1 + \tilde{\lambda}}}{s + (1 + \tilde{\lambda})}.
\end{aligned}$$

Also, for $L' \neq L$,

$$\begin{aligned}
\tilde{p}(F_{L'M}^{A\bar{A}} b|E \cap F_L^A a) &= \sum_{M'} \tilde{p}(F_{L'M}^{A\bar{A}} b|E \cap F_{LM'}^{A\bar{A}} a)\xi_{M'} \\
&= \frac{s_{L'M}^{A\bar{A}} + \tilde{\lambda}\tilde{\gamma}_{L'M}^{A\bar{A}}}{s + 1 + \tilde{\lambda}} \sum_{M'} \xi_{M'}
\end{aligned}$$

$$= \frac{s^{A\bar{A}}_{L'M} + \tilde{\lambda}\tilde{\gamma}^{A\bar{A}}_{L'M}}{s + 1 + \tilde{\lambda}}$$

$$= \frac{s^{A\bar{A}}_{L'M} + (1 + \tilde{\lambda})\frac{\tilde{\lambda}\tilde{\gamma}^{A\bar{A}}_{L'M}}{1+\tilde{\lambda}}}{s + (1 + \tilde{\lambda})}.$$

Hence $\mathcal{F}^{1\ldots n}$ is a λ-family relative to $\tilde{p}(\cdot|F^A_L a)$ for individuals not involved in D or equal to a. Thus the lemma holds for $t = \tau + 1$ and so by mathematical induction it holds for all t.

The next lemma is a generalization of Theorem 10 of Maher (2000).

LEMMA 9. Let $p = \sum_{\mathcal{A}} c^{\mathcal{A}} p^{\mathcal{A}}$, where the summation is over all partitions $\mathcal{A}$ of $\{1, \ldots, n\}$ and the $c^{\mathcal{A}}$ are non-negative constants that sum to 1, with $c^{1\ldots n} > 0$. Let $\{E_s\}_{s=1}^{\infty}$ be a sequence of sample propositions with respect to $\mathcal{F}^{1\ldots n}$, where each E_s is for a sample of size s and E_{s+1} entails E_s. Let $s^{1\ldots n}_{l_1\ldots l_n}$ denote the number of individuals that are ascribed $F^{1\ldots n}_{l_1\ldots l_n}$ by E_s. Then for any individual a not involved in any of the E_s,

$$\lim_{s\to\infty} \left| p(F^{1\ldots n}_{l_1\ldots l_n} a|E_s) - \frac{s^{1\ldots n}_{l_1\ldots l_n}}{s} \right| = 0.$$

Proof. It follows from Definition 5 that each $p^{\mathcal{A}}$ has the following property: If E and E' are sample propositions with respect to $\mathcal{F}^{1\ldots n}$ that ascribe each property in $\mathcal{F}^{1\ldots n}$ to the same number of individuals then $p^{\mathcal{A}}(E) = p^{\mathcal{A}}(E')$. Since $p = \sum_{\mathcal{A}} c^{\mathcal{A}} p^{\mathcal{A}}$, it follows that p also has this property.

Let $N = 2^n$ and let the N properties in $\mathcal{F}^{1\ldots n}$ be enumerated in some way. Given a sample proposition E with respect to $\mathcal{F}^{1\ldots n}$, let s_i denote the number of individuals to which E ascribes the property that has the ith place in this enumeration. Let

$$S = \{(x_1, \ldots, x_{N-1}) \colon x_i > 0, \textstyle\sum_{i=1}^{N-1} x_i < 1\}.$$

Thus S is a simplex of dimension $N - 1$. Let $x_N = 1 - \sum_{i=1}^{N-1} x_i$. By de Finetti's representation theorem there exists on S a unique probability measure μ and, for each $\mathcal{A}$, a unique probability measure $\mu_{\mathcal{A}}$, such that for any sample proposition E with respect to $\mathcal{F}^{1\ldots n}$,

$$(16) \qquad p(E) = \int_S \prod_{i=1}^{N} x_i^{s_i}\, d\mu(x_1, \ldots, x_{N-1}),$$

$$p^{\mathcal{A}}(E) = \int_S \prod_{i=1}^{N} x_i^{s_i}\, d\mu^{\mathcal{A}}(x_1, \ldots, x_{N-1}).$$

Now $\sum_{\mathcal{A}} c^{\mathcal{A}} \mu^{\mathcal{A}}$ is also a probability measure on S and (suppressing the variables of integration for clarity) we have

$$\begin{aligned} \int_S \prod_{i=1}^{N} x_i^{s_i}\, d(\textstyle\sum_{\mathcal{A}} c^{\mathcal{A}} \mu^{\mathcal{A}}) &= \sum_{\mathcal{A}} c^{\mathcal{A}} \int_S \prod_{i=1}^{N} x_i^{s_i}\, d\mu^{\mathcal{A}} \\ &= \sum_{\mathcal{A}} c^{\mathcal{A}} p^{\mathcal{A}}(E) && \text{by (16)} \\ &= p(E). \end{aligned}$$

So $\mu = \sum_{\mathcal{A}} c^{\mathcal{A}} \mu^{\mathcal{A}}$. Since $\mathcal{F}^{1\ldots n}$ is a λ-family with respect to $p^{1\ldots n}$, $\mu^{1\ldots n}$ is a Dirichlet distribution on S (Festa 1993, §6.3). It follows that $\mu^{1\ldots n}(B) > 0$ for all open non-empty $B \subset S$. Thus, for any such B,

$$\mu(B) = \sum_{\mathcal{A}} c^{\mathcal{A}} \mu^{\mathcal{A}}(B) \geq c^{1\ldots n} \mu^{1\ldots n}(B) > 0.$$

Hence the support of μ (the set of points for which all open neighborhoods have positive measure with respect to μ) is the whole of S. Lemma 9 now follows from Lemma 8 of Fine (1973, 194).

The next lemma generalizes the preceding one.

LEMMA 10. Let $p = \sum_{\mathcal{A}} c^{\mathcal{A}} p^{\mathcal{A}}$, where the summation is over all partitions $\mathcal{A}$ of $\{1, \ldots, n\}$ and the $c^{\mathcal{A}}$ are non-negative constants that sum to 1, with $c^{1\ldots n} > 0$. Let $B \subset \{1, \ldots, n\}$, $B \neq \emptyset$. Let $\{E^B_{(s)}\}_{s=1}^{\infty}$ be a sequence of sample propositions with respect to $\mathcal{F}^B$, where each $E^B_{(s)}$ is for a sample of size s and $E^B_{(s+1)}$ entails $E^B_{(s)}$. Let s^B_L denote the number of individuals that are ascribed F^B_L by $E^B_{(s)}$. Then for any individual a not involved in any of the $E^B_{(s)}$,

$$\lim_{s\to\infty} \left| p(F^B_L a | E^B_{(s)}) - \frac{s^B_L}{s} \right| = 0.$$

Proof. Let D^B be a sample proposition with respect to $\mathcal{F}^B$ and let D be any sample proposition with respect to $\mathcal{F}^{1\ldots n}$ that involves the same individuals as D and is such that $D \subset D^B$. Then

$$\begin{aligned}
p^B(D^B) &= p(D^B) && \text{by Definition 5}\\
&= \sum_D p(D)\\
&= \sum_D \sum_{\mathcal{A}} c^{\mathcal{A}} p^{\mathcal{A}}(D)\\
&= \sum_{\mathcal{A}} c^{\mathcal{A}} p^{\mathcal{A}}(D^B)\\
&= \sum_{\mathcal{A}} c^{\mathcal{A}} p^{\mathcal{A}_B}(D^B) && \text{by Lemma 2.}
\end{aligned}$$

For any partition $\mathcal{B}$ of B let $c^{\mathcal{B}} \stackrel{\text{df}}{=} \sum_{\mathcal{A}_B=\mathcal{B}} c^{\mathcal{A}}$. Then the preceding equation can be rewritten as

$$p^B(D^B) = \sum_{\mathcal{B}} c^{\mathcal{B}} p^{\mathcal{B}}(D^B).$$

Now applying Lemma 9 with B in place of $\{1, \ldots, n\}$ gives

$$\lim_{s\to\infty} \left| p^B(F_L^B a | E_{(s)}^B) - \frac{s_L^B}{s} \right| = 0.$$

Since p agrees with p^B on $\mathcal{X}^B$ it follows that

$$\lim_{s\to\infty} \left| p(F_L^B a | E_{(s)}^B) - \frac{s_L^B}{s} \right| = 0.$$

8.6. *Proof of Theorem 4*

I will prove the theorem for the special case in which $n = 3$. It follows by Lemma 2 that the theorem also holds for $n > 3$.

By Lemma 7, $p(F_{11}^{12} a | E_{(s)}^{12})$ equals

$$p^{1|2|3}(F_{11}^{12} a | E_{(s)}^{12}) p(I^{1|2|3} | E_{(s)}^{12}) + p^{123}(F_{11}^{12} a | E_{(s)}^{12}) p(I^{123} | E_{(s)}^{12}).$$

Let $r_L^A = s_L^A / s$. It follows that

$$p(F_{11}^{12} a | E_{(s)}^{12}) - r_{11}^{12} = \left[p^{1|2|3}(F_{11}^{12} a | E_{(s)}^{12}) - r_{11}^{12}\right] p(I^{1|2|3} | E_{(s)}^{12}) +$$
$$\left[p^{123}(F_{11}^{12} a | E_{(s)}^{12}) - r_{11}^{12}\right] p(I^{123} | E_{(s)}^{12}).$$

Lemma 10 entails

$$\lim_{s\to\infty} \left| p(F^{12}_{11}a|E^{12}_{(s)}) - r^{12}_{11} \right| = 0$$

and

$$\lim_{s\to\infty} \left| p^{123}(F^{12}_{11}a|E^{12}_{(s)}) - r^{12}_{11} \right| = 0.$$

Hence

$$\text{(17)} \qquad \lim_{s\to\infty} \left| p^{1|2|3}(F^{12}_{11}a|E^{12}_{(s)}) - r^{12}_{11} \right| p(I^{1|2|3}|E^{12}_{(s)}) = 0.$$

But

$$\begin{aligned} \lim_{s\to\infty} \left| p^{1|2|3}(F^{12}_{11}a|E^{12}_{(s)}) - r^{12}_{11} \right| &= \lim_{s\to\infty} \left| p^{1|2}(F^{12}_{11}a|E^{12}_{(s)}) - r^{12}_{11} \right|, \\ &\qquad \text{by Lemma 2} \\ &= \lim_{s\to\infty} \left| \frac{s^1_1 + \lambda\gamma^1_1}{s+\lambda} \frac{s^2_1 + \lambda\gamma^2_1}{s+\lambda} - r^{12}_{11} \right| \\ &= \lim_{s\to\infty} \left| r^1_1 r^2_1 - r^{12}_{11} \right| \end{aligned}$$

and

$$\begin{aligned} \left| r^1_1 r^2_1 - r^{12}_{11} \right| &= r^2_1 r^2_2 \left| \frac{r^1_1 r^2_1 - r^{12}_{11}}{r^2_1 r^2_2} \right| \\ &= r^2_1 r^2_2 \left| \frac{r^1_1 r^2_1 - r^2_1 r^{12}_{11} - r^{12}_{11} + r^2_1 r^{12}_{11}}{r^2_1 r^2_2} \right| \\ &= r^2_1 r^2_2 \left| \frac{r^2_1 r^{12}_{12} - r^2_2 r^{12}_{11}}{r^2_1 r^2_2} \right| \\ &= r^2_1 r^2_2 \left| \frac{r^{12}_{12}}{r^2_2} - \frac{r^{12}_{11}}{r^2_1} \right| \\ &= \frac{s^2_1}{s} \frac{s^2_2}{s} \left| \frac{s^{12}_{12}}{s^2_2} - \frac{s^{12}_{11}}{s^2_1} \right| \\ &> \varepsilon^3 \qquad\qquad \text{for all } s > S. \end{aligned}$$

Hence

$$\lim_{s\to\infty} \left| p^{1|2|3}(F^{12}_{11}a|E^{12}_{(s)}) - r^{12}_{11} \right| \neq 0.$$

By (17) it follows that

$$\text{(18)} \qquad \lim_{s\to\infty} p(I^{1|2|3}|E^{12}_{(s)}) = 0.$$

For $l = 1$ or 2,

$$\lim_{s\to\infty} p(F_1^3 a | F_{11}^{12} a \cap F_{12l}^{123} b \cap E_{(s)}^{12}) =$$

$$= \lim_{s\to\infty} \frac{p(F_{111}^{123} a \cap F_{12l}^{123} b | E_{(s)}^{12})}{p(F_{11}^{12} a \cap F_{12l}^{123} b | E_{(s)}^{12})}$$

$$= \lim_{s\to\infty} \frac{p^{123}(F_{111}^{123} a \cap F_{12l}^{123} b | E_{(s)}^{12})}{p^{123}(F_{11}^{12} a \cap F_{12l}^{123} b | E_{(s)}^{12})} \text{ by Lemma 7 and (18)}$$

$$= \lim_{s\to\infty} p^{123}(F_1^3 a | F_{11}^{12} a \cap F_{12l}^{123} b \cap E_{(s)}^{12}).$$

By Lemma 8, $\mathcal{F}^{123}$ is a λ-family with respect to $p^{123}(\cdot | E_{(s)}^{12})$ for individuals not involved in $E_{(s)}^{12}$. So by the argument given in Section 3,

$$p^{123}(F_1^3 a | F_{11}^{12} a \cap F_{12l}^{123} b \cap E_{(s)}^{12})$$

is independent of l. Hence

$$\lim_{s\to\infty} p(F_1^3 a | F_{11}^{12} a \cap F_{12l}^{123} b \cap E_{(s)}^{12})$$

is independent of l. The theorem is an immediate consequence of this.

8.7. *Lemmas Used in the Proof of Theorem 5*

LEMMA 11. Let $A \subset \{1, \ldots, n\}$, $A \neq \emptyset$, $A \neq \{1, \ldots, n\}$. Let $\mathcal{F}^{1\ldots n}$ be a λ-family relative to p with $\gamma_{LM}^{A\bar{A}} = \gamma_L^A \gamma_M^{\bar{A}}$. Let E^A be a sample proposition with respect to $\mathcal{F}^A$. By Lemma 8, $\mathcal{F}^{1\ldots n}$ is a λ-family relative to $p(\cdot | E^A)$ for individuals not involved in E^A; let $\tilde{\lambda}$ and $\tilde{\gamma}$ denote parameters for $p(\cdot | E^A)$. Then $\tilde{\gamma}_{LM}^{A\bar{A}} = \tilde{\gamma}_L^A \tilde{\gamma}_M^{\bar{A}}$.

Proof. Let s denote the number of individuals involved in E^A. If $s = 0$ then the lemma is trivially true. Now suppose the lemma holds for $s = \sigma \geq 0$. Let $\tilde{p} = p(\cdot | E^A)$ and let $\tilde{\lambda}$ and $\tilde{\gamma}$ denote parameters for $\tilde{p}$. Let a and b be distinct individuals not involved in E^A. Let $F_L^A \in \mathcal{F}^A$ and let $\hat{\lambda}$ and $\hat{\gamma}$ denote parameters for $\tilde{p}(\cdot | F_L^A a)$. Then for any M,

$$\text{(19)} \qquad \tilde{p}(F_M^{\bar{A}} a | F_L^A a) = \frac{\tilde{p}(F_{LM}^{A\bar{A}} a)}{\tilde{p}(F_L^A a)}$$

$$= \frac{\tilde{\gamma}^{A\bar{A}}_{LM}}{\tilde{\gamma}^{A}_{L}}$$

$$= \tilde{\gamma}^{\bar{A}}_{M} \qquad \text{by assumption.}$$

$$\begin{aligned}
(20)\quad \hat{\gamma}^{\bar{A}}_{M} &= \tilde{p}(F^{\bar{A}}_{M}b|F^{A}_{L}a) \\
&= \sum_{M'} \tilde{p}(F^{\bar{A}}_{M}b|F^{A\bar{A}}_{LM'}a)\tilde{p}(F^{\bar{A}}_{M'}a|F^{A}_{L}a) \\
&= \sum_{M'} \tilde{p}(F^{\bar{A}}_{M}b|F^{A\bar{A}}_{LM'}a)\tilde{\gamma}^{\bar{A}}_{M'} && \text{by (19)} \\
&= \sum_{M'}\sum_{L'} \tilde{p}(F^{A\bar{A}}_{L'M}b|F^{A\bar{A}}_{LM'}a)\tilde{\gamma}^{\bar{A}}_{M'} \\
&= \sum_{L'} \tilde{p}(F^{A\bar{A}}_{L'M}b|F^{A\bar{A}}_{LM}a)\tilde{\gamma}^{\bar{A}}_{M} + \sum_{M'\neq M}\sum_{L'} \tilde{p}(F^{A\bar{A}}_{L'M}b|F^{A\bar{A}}_{LM}a)\tilde{\gamma}^{\bar{A}}_{M'} \\
&= \frac{1+\sum_{L'}\tilde{\lambda}\tilde{\gamma}^{A\bar{A}}_{L'M}}{1+\tilde{\lambda}}\tilde{\gamma}^{\bar{A}}_{M} + \sum_{M'\neq M}\frac{\sum_{L'}\tilde{\lambda}\tilde{\gamma}^{A\bar{A}}_{L'M}}{1+\tilde{\lambda}}\tilde{\gamma}^{\bar{A}}_{M'} \\
&= \frac{1+\tilde{\lambda}\tilde{\gamma}^{\bar{A}}_{M}}{1+\tilde{\lambda}}\tilde{\gamma}^{\bar{A}}_{M} + \sum_{M'\neq M}\frac{\tilde{\lambda}\tilde{\gamma}^{\bar{A}}_{M}}{1+\tilde{\lambda}}\tilde{\gamma}^{\bar{A}}_{M'} \\
&= \tilde{\gamma}^{\bar{A}}_{M}.
\end{aligned}$$

$$\begin{aligned}
(21)\quad \hat{\gamma}^{A}_{L} &= \tilde{p}(F^{A}_{L}b|F^{A}_{L}a) \\
&= \frac{1+\tilde{\lambda}\tilde{\gamma}^{A}_{L}}{1+\tilde{\lambda}} && \text{by Theorem 1.}
\end{aligned}$$

$$\begin{aligned}
(22)\quad \hat{\gamma}^{A\bar{A}}_{LM} &= \tilde{p}(F^{A\bar{A}}_{LM}b|F^{A}_{L}a) \\
&= \sum_{M'} \tilde{p}(F^{A\bar{A}}_{LM}b|F^{A\bar{A}}_{LM'}a)\tilde{p}(F^{\bar{A}}_{M'}a|F^{A}_{L}a) \\
&= \sum_{M'} \tilde{p}(F^{A\bar{A}}_{LM}b|F^{A\bar{A}}_{LM'}a)\tilde{\gamma}^{\bar{A}}_{M'} && \text{by (19)} \\
&= \frac{1+\tilde{\lambda}\tilde{\gamma}^{A\bar{A}}_{LM}}{1+\tilde{\lambda}}\tilde{\gamma}^{\bar{A}}_{M} + \sum_{M'\neq M}\frac{\tilde{\lambda}\tilde{\gamma}^{A\bar{A}}_{LM}}{1+\tilde{\lambda}}\tilde{\gamma}^{\bar{A}}_{M'} \\
&= \frac{\tilde{\gamma}^{\bar{A}}_{M}+\tilde{\lambda}\tilde{\gamma}^{A\bar{A}}_{LM}}{1+\tilde{\lambda}}
\end{aligned}$$

$$= \tilde{\gamma}_M^{\bar{A}} \frac{1 + \tilde{\lambda}\tilde{\gamma}_L^A}{1 + \tilde{\lambda}}$$

$$= \hat{\gamma}_L^A \hat{\gamma}_M^{\bar{A}} \qquad \text{by (20) and (21).}$$

If $L' \neq L$ then

$$\text{(23)} \quad \hat{\gamma}_{L'}^A = \tilde{p}(F_{L'}^A b | F_L^A a) = \frac{\tilde{\lambda}\tilde{\gamma}_{L'}^A}{1 + \tilde{\lambda}} \qquad \text{by Theorem 1.}$$

So for $L' \neq L$,

$$\begin{aligned}
\text{(24)} \quad \hat{\gamma}_{L'M}^{A\bar{A}} &= \tilde{p}(F_{L'M}^{A\bar{A}} b | F_L^A a) \\
&= \sum_{M'} \tilde{p}(F_{L'M}^{A\bar{A}} b | F_{LM'}^{A\bar{A}} a) \tilde{p}(F_{M'}^{\bar{A}} a | F_L^A a) \\
&= \sum_{M'} \tilde{p}(F_{L'M}^{A\bar{A}} b | F_{LM'}^{A\bar{A}} a) \tilde{\gamma}_{M'}^{\bar{A}} && \text{by (19)} \\
&= \sum_{M'} \frac{\tilde{\lambda}\tilde{\gamma}_{L'M}^{A\bar{A}}}{1 + \tilde{\lambda}} \tilde{\gamma}_{M'}^{\bar{A}} \\
&= \frac{\tilde{\lambda}\tilde{\gamma}_{L'M}^{A\bar{A}}}{1 + \tilde{\lambda}} \\
&= \frac{\tilde{\lambda}\tilde{\gamma}_{L'}^A}{1 + \tilde{\lambda}} \tilde{\gamma}_M^{\bar{A}} && \text{by assumption} \\
&= \hat{\gamma}_{L'}^A \hat{\gamma}_M^{\bar{A}} && \text{by (20) and (23).}
\end{aligned}$$

Together (22) and (24) show that the lemma holds for $s = \sigma + 1$. So by mathematical induction the lemma holds for all s.

LEMMA 12. Let $A \subset \{1, \ldots, n\}$, $A \neq \emptyset$, $A \neq \{1, \ldots, n\}$. Let E^A be a sample proposition with respect to $\mathcal{F}^A$; by Lemma 8, $\mathcal{F}^{1\ldots n}$ is a λ-family relative to $p^{1\ldots n}(\cdot|E^A)$ for individuals not involved in E^A. Then $\mathcal{F}^A$ and $\mathcal{F}^{\bar{A}}$ are probabilistically independent λ-families relative to $p^{A|\bar{A}}(\cdot|E^A)$ for individuals not involved in E^A and have the same λ and γ values as in $p^{1\ldots n}(\cdot|E^A)$.

Proof. Let D be a sample proposition with respect to $\mathcal{F}^{1\ldots n}$. Let D^A be the sample proposition with respect to $\mathcal{F}^A$ that involves the same individuals as D and is entailed by D; similarly for $D^{\bar{A}}$. Let $\tilde{D}^A$ denote an arbitrary sample proposition with respect to $\mathcal{F}^A$ that involves the same individuals as D; similarly for $\tilde{E}^{\bar{A}}$.

$$
\text{(25)}\quad \begin{aligned}
p^{A|\bar{A}}(D^{\bar{A}}|E^A) &= \frac{p^{A|\bar{A}}(D^{\bar{A}} \cap E^A)}{p^{A|\bar{A}}(E^A)} \\
&= \sum_{\tilde{D}^A} \sum_{\tilde{E}^{\bar{A}}} \frac{p^{A|\bar{A}}(\tilde{D}^A \cap D^{\bar{A}} \cap E^A \cap \tilde{E}^{\bar{A}})}{p^{A|\bar{A}}(E^A)} \\
&= \sum_{\tilde{D}^A} \sum_{\tilde{E}^{\bar{A}}} \frac{p^{A|\bar{A}}(\tilde{D}^A \cap E^A) p^{A|\bar{A}}(D^{\bar{A}} \cap \tilde{E}^{\bar{A}})}{p^{A|\bar{A}}(E^A)} \\
&= \frac{p^{A|\bar{A}}(E^A) p^{A|\bar{A}}(D^{\bar{A}})}{p^{A|\bar{A}}(E^A)} \\
&= p^{A|\bar{A}}(D^{\bar{A}}).
\end{aligned}
$$

$$
\begin{aligned}
p^{A|\bar{A}}(D|E^A) &= \frac{p^{A|\bar{A}}(D \cap E^A)}{p^{A|\bar{A}}(E^A)} \\
&= \sum_{\tilde{E}^{\bar{A}}} \frac{p^{A|\bar{A}}(D \cap E^A \cap \tilde{E}^{\bar{A}})}{p^{A|\bar{A}}(E^A)} \\
&= \sum_{\tilde{E}^{\bar{A}}} \frac{p^{A|\bar{A}}(D^A \cap E^A) p^{A|\bar{A}}(D^{\bar{A}} \cap \tilde{E}^{\bar{A}})}{p^{A|\bar{A}}(E^A)} \\
&= \frac{p^{A|\bar{A}}(D^A \cap E^A) p^{A|\bar{A}}(D^{\bar{A}})}{p^{A|\bar{A}}(E^A)} \\
&= p^{A|\bar{A}}(D^A|E^A) p^{A|\bar{A}}(D^{\bar{A}}) \\
&= p^{A|\bar{A}}(D^A|E^A) p^{A|\bar{A}}(D^{\bar{A}}|E^A) \qquad \text{by (25).}
\end{aligned}
$$

So by Definition 2, $\mathcal{F}^A$ and $\mathcal{F}^{\bar{A}}$ are probabilistically independent relative to $p^{A|\bar{A}}(\cdot|E^A)$.

Now let a be an individual not involved in E^A. Let λ and γ denote parameter values for $p^{1\ldots n}(\cdot|E^A)$. Then

$$
\begin{aligned}
p^{A|\bar{A}}(F_L^A a|E^A) &= p^{1\ldots n}(F_L^A a|E^A) \qquad \text{by Lemma 2,} \\
p^{A|\bar{A}}(F_M^{\bar{A}}|E^A) &= p^{A|\bar{A}}(F_M^{\bar{A}} a) \qquad \text{by (25)} \\
&= p^{1\ldots n}(F_M^{\bar{A}} a) \qquad \text{by Lemma 2} \\
&= \sum_{\tilde{E}^{\bar{A}}} p^{1\ldots n}(F_M^{\bar{A}} a|\tilde{E}^{\bar{A}}) p^{1\ldots n}(\tilde{E}^{\bar{A}})
\end{aligned}
$$

$$= \sum_{\tilde{E}^{\bar{A}}} p^{1\ldots n}(F_M^{\bar{A}}a|E^A \cap \tilde{E}^{\bar{A}})p^{1\ldots n}(\tilde{E}^{\bar{A}}|E^A), \quad \text{by Lemma 4}$$

$$= p^{1\ldots n}(F_M^{\bar{A}}a|E^A).$$

Thus $\mathcal{F}^A$ and $\mathcal{F}^{\bar{A}}$ have the same λ and γ values in $p^{A|\bar{A}}$ as in $p^{1\ldots n}$.

8.8. *Proof of Theorem 5*

I will prove the theorem for the special case in which $n = 3$. If follows by Lemma 2 that the theorem also holds for $n > 3$.

Let $r_L^A = s_L^A/s$. By Lemma 6,

$$p(F_{11}^{23}a|E_{(s)}^{23}) - r_{11}^{23} = \sum_{\mathcal{A}} \left|p^{\mathcal{A}}(F_{11}^{23}a|E_{(s)}^{23}) - r_{11}^{23}\right| p(I^{\mathcal{A}}|E_{(s)}^{23}).$$

If $\tilde{p}$ is p, $p^{1|23}$, or p^{123}, then by Lemma 10,

$$\lim_{s\to\infty} \left|\tilde{p}(F_{11}^{23}a|E_{(s)}^{23}) - r_{11}^{23}\right| = 0.$$

Hence

$$\begin{aligned}
&\lim\nolimits_{s\to\infty} \left|p^{1|2|3}(F_{11}^{23}a|E_{(s)}^{23}) - r_{11}^{23}\right| p(I^{1|2|3}|E_{(s)}^{23}) + \\
&\lim\nolimits_{s\to\infty} \left|p^{12|3}(F_{11}^{23}a|E_{(s)}^{23}) - r_{11}^{23}\right| p(I^{12|3}|E_{(s)}^{23}) + \\
&\lim\nolimits_{s\to\infty} \left|p^{2|13}(F_{11}^{23}a|E_{(s)}^{23}) - r_{11}^{23}\right| p(I^{2|13}|E_{(s)}^{23}) \qquad = 0.
\end{aligned}$$

Hence, by Lemma 2,

$$\begin{aligned}
(26) \qquad &\lim_{s\to\infty} \left|p^{2|3}(F_{11}^{23}a|E_{(s)}^{23}) - r_{11}^{23}\right| \cdot \\
&[p(I^{1|2|3}|E_{(s)}^{23}) + p(I^{12|3}|E_{(s)}^{23}) + p(I^{2|13}|E_{(s)}^{23})] = 0.
\end{aligned}$$

But

$$\begin{aligned}
\lim_{s\to\infty} \left|p^{2|3}(F_{11}^{23}a|E_{(s)}^{23}) - r_{11}^{23}\right| &= \lim_{s\to\infty} \left|\frac{s_1^2 + \lambda\gamma_1^2}{s+\lambda}\frac{s_1^3 + \lambda\gamma_1^3}{s+\lambda} - r_{11}^{23}\right| \\
&= \lim_{s\to\infty} \left|r_1^2 r_1^3 - r_{11}^{23}\right|
\end{aligned}$$

and

$$\begin{aligned}
\left|r_1^2 r_1^3 - r_{11}^{23}\right| &= r_1^2 r_2^2 \left|\frac{r_1^2 r_1^3 - r_{11}^{23}}{r_1^2 r_2^2}\right| \\
&= r_1^2 r_2^2 \left|\frac{r_1^2 r_1^3 - r_1^2 r_{11}^{23} - r_{11}^{23} + r_1^2 r_{11}^{23}}{r_1^2 r_2^2}\right| \\
&= r_1^2 r_2^2 \left|\frac{r_1^2 r_{21}^{23} - r_2^2 r_{11}^{23}}{r_1^2 r_2^2}\right| \\
&= r_1^2 r_2^2 \left|\frac{r_{21}^{23}}{r_2^2} - \frac{r_{11}^{23}}{r_1^2}\right| \\
&= \frac{s_1^2}{s}\frac{s_2^2}{s}\left|\frac{s_{21}^{23}}{s_2^2} - \frac{s_{11}^{23}}{s_1^2}\right| \\
&> \varepsilon^3 \qquad\qquad \text{for all } s > S.
\end{aligned}$$

Thus

$$\lim_{s\to\infty}\left|p^{2|3}(F_{11}^{23}a|E_{(s)}^{23}) - r_{11}^{23}\right| \neq 0.$$

By (26) it follows that

$$\text{(27)} \qquad \lim_{s\to\infty}[p(I^{1|2|3}|E_{(s)}^{23}) + p(I^{12|3}|E_{(s)}^{23}) + p(I^{2|13}|E_{(s)}^{23})] = 0.$$

For $l = 1$ or 2,

$$\begin{aligned}
&\lim_{s\to\infty} p(F_1^3 a|F_{11}^{12}a \cap F_{12l}^{123}b \cap E_{(s)}^{23}) \\
&= \lim_{s\to\infty}\frac{p(F_{111}^{123}a \cap F_{12l}^{123}b|E_{(s)}^{23})}{p(F_{11}^{12}a \cap F_{12l}^{123}b|E_{(s)}^{23})} \\
&= \lim_{s\to\infty}\big\{\big[p^{123}(F_{111}^{123}a \cap F_{12l}^{123}b|E_{(s)}^{23})p(I^{123}|E_{(s)}^{23}) + \\
&\quad p^{1|23}(F_{111}^{123}a \cap F_{12l}^{123}b|E_{(s)}^{23})p(I^{1|23}|E_{(s)}^{23})\big] / \\
&\quad \big[p^{123}(F_{11}^{12}a \cap F_{12l}^{123}b|E_{(s)}^{23})p(I^{123}|E_{(s)}^{23}) + \\
&\quad p^{1|23}(F_{11}^{12}a \cap F_{12l}^{123}b|E_{(s)}^{23})p(I^{1|23}|E_{(s)}^{23})\big]\big\} \qquad \text{by (27) and Lemma 6.}
\end{aligned}$$

By Lemma 8, $\mathcal{F}^{123}$ is a λ-family with respect to $p^{123}(\cdot|E_{(s)}^{23})$. By Lemma 12, $\mathcal{F}^{1}$ and $\mathcal{F}^{23}$ are probabilistically independent λ-families with respect to $p^{1|23}(\cdot|E^{23}_{(s)})$ and have the same λ and γ values as in $p^{123}(\cdot|E^{23}_{(s)})$. Letting λ and γ denote these common values, we have by Definition 1 that the last expression is equal to

$$\lim_{s\to\infty} \frac{\gamma^{123}_{12l}\frac{\lambda\gamma^{123}_{111}}{1+\lambda}p(I^{123}|E^{23}_{(s)}) + \gamma^1_1\frac{1+\lambda\gamma^1_1}{1+\lambda}\gamma^{23}_{2l}\frac{\lambda\gamma^{23}_{11}}{1+\lambda}p(I^{1|23}|E^{23}_{(s)})}{\gamma^{123}_{12l}\frac{\lambda\gamma^{12}_{11}}{1+\lambda}p(I^{123}|E^{23}_{(s)}) + \gamma^1_1\frac{1+\lambda\gamma^1_1}{1+\lambda}\gamma^{23}_{2l}\frac{\lambda\gamma^2_1}{1+\lambda}p(I^{1|23}|E^{23}_{(s)})}.$$

By Lemma 11, $\gamma^{123}_{l_1l_2l_3} = \gamma^1_{l_1}\gamma^{23}_{l_2l_3}$. So dividing numerator and denominator by γ^{123}_{12l} gives an expression in which l does not appear. Hence

$$\lim_{s\to\infty} p(F^3_1 a|F^{12}_{11}a \cap F^{123}_{12l}b \cap E^{23}_{(s)})$$

is independent of l. The theorem is an immediate consequence of this.

ACKNOWLEDGEMENTS

I thank Roberto Festa and Theo Kuipers for comments that helped improve this paper.

NOTES

1 In many treatments of analogy in the literature, properties of the form $F^{1\ldots n}_{l_1\ldots l_n}$ are denoted $Q_1, \ldots, Q_k$, where $k = 2^n$, and other properties are represented as disjunctions of these "Q-properties". I used that notation myself in (Maher, 2000), where I dealt with the case $n = 2$. However, many of the results of this paper – beginning with Theorem 1 – cannot feasibly be expressed in terms of Q-properties, so I will not use the Q-property notation in this paper. It may also be worth noting that even the results of (Maher 2000) can often be stated and proved more efficiently using the notation of this paper. For example, the four equations of Theorem 5 of (Maher 2000) can be expressed in the present notation with the single equation

$$p(F^{12}_{lm}a|E \cap I) = \frac{s^1_l + \lambda\gamma^1_l}{s+\lambda}\frac{s^2_m + \lambda\gamma^2_m}{s+\lambda}.$$

2 The condition stated here accords with the traditional conception of analogical reasoning that can be found in Hume (1748, §82), Mill (1874, bk. III, ch. XX, §2), Keynes (1921, ch. XIX), Carnap (1945), Achinstein (1963), Hesse (1964), and introductory logic texts such as Copi and Cohen (1998). A different condition, analogous to Carnap's axiom of analogy CA that I discussed in (Maher 2000), would be:

$$p(F^{123}_{111}a|F^{123}_{121}b) > p(F^{123}_{111}a|F^{123}_{122}b).$$

Kuipers (1984, 73) advocated another analogy condition that is like this one in relating Q-properties. I will not discuss conditions of the latter kind in the present paper because I think they are less intuitively compelling than the one stated in the text. However, it is easy to show that P_λ does not satisfy these other analogy conditions and I believe – though I have not proved – that analogs of the main negative results of this paper (Theorems 4 and 5) also hold for these other analogy conditions.

REFERENCES

Achinstein, P.: 1963, 'Variety and Analogy in Confirmation Theory', *Philosophy of Science* **30**, 207–221.

Carnap, R.: 1945, 'On Inductive Logic', *Philosophy of Science* **12**, 72–97.

Carnap, R.: 1952, *The Continuum of Inductive Methods*, University of Chicago Press, Chicago.

Carnap, R.: 1954, 'm(Z) for n Families as a Combination of m_λ-Functions', Document 093–22–01 in the Rudolf Carnap Collection, University of Pittsburgh Library. The notation is explained in document 093–22-03.

Carnap, R.: 1963, 'Replies and Systematic Expositions', in P. A. Schilpp (ed.), *The Philosophy of Rudolf Carnap*, Open Court, La Salle, IL, pp. 859–1013.

Carnap, R.: 1975, 'Notes on Probability and Induction', in J. Hintikka (ed.), *Rudolf Carnap, Logical Empiricist*, Reidel, Dordrecht, pp. 293–324.

Carnap, R. and W. Stegmüller: 1959, *Induktive Logik und Wahrscheinlichkeit*, Springer, Wien.

Copi, I. M. and C. Cohen: 1998, *Introduction to Logic*, Prentice Hall, Upper Saddle River, NJ, 10th edn.

Festa, R.: 1993, *Optimum Inductive Methods*, Kluwer, Dordrecht.

Fine, T. L.: 1973, *Theories of Probability*, Academic Press, New York.

Hesse, M.: 1964, 'Analogy and Confirmation Theory', *Philosophy of Science* **31**, 319–327.

Hume, D.: 1748, *An Enquiry Concerning Human Understanding*. many reprints.

Keynes, J. M.: 1921, *A Treatise on Probability*, Macmillan, London.

Kuipers, T. A. F.: 1984, 'Two Types of Inductive Analogy by Similarity', *Erkenntnis* **21**, 63–87.

Maher, P.: 2000, 'Probabilities for Two Properties', *Erkenntnis* **52**, 63–91.

Mill, J. S.: 1874, *A System of Logic*, Harper, New York, 8th edition.

Department of Philosophy
University of Illinois
105 Gregory Hall
810 S. Wright Street
Urbana IL 61801
U.S.A.
E-mail: p-maher@uiuc.edu