Binary Induction and Carnap's Continuum

Alena Vencovska*
School of Mathematics
The University of Manchester, UK
alena.vencovska@manchester.ac.uk

Abstract

We consider the problem of induction over languages with binary predicates and show that a natural generalization of Johnson's Sufficientness Postulate eliminates all but two solutions. We discuss the historical context and connections to the unary case.

1 Introduction

In his posthumously published paper of 1932 [1], W.E.Johnson introduced a principle of inductive reasoning, subsequently called *Johnson's Sufficientness Principle* by I.J.Good, which was to be independently rediscovered some 20 years later by Rudolf Carnap in his programme of formulating what he termed *Inductive Logic*. Roughly (a precise statement will be given later) the principle asserts that if we run an experiment with some finite number of outcomes, $b_1, b_2, \ldots, b_k$ say, some n times then the probability we should give to the $n+1$st run of the experiment yielding a particular outcome b_1 should only be a function of n and the number of previous occasions on which the outcome b_1 had been observed.

This was a central principle for Johnson and Carnap because imposing it, along with some few other generally uncontentious requirements, constrained the assigned probabilities to lie within a particularly simple parameterized family now referred to as Carnap's Continuum of Inductive Methods.

Despite this remarkable success of the principle it is certainly not without criticism (for an extensive discussion see chapter 4 of [6]). For example should not the distribution of outcomes amongst the other possible $b_2, b_3, \ldots, b_k$ also be relevant to the probability one should assign to b_1 on the $n+1$st run? Even the advantage that this principle yielded a simple parameterized family of possibilities has since been challenged by the alternative *Nix-Paris Continuum*,

*Supported by a UK Engineering and Physical Sciences Research Council (EPSRC) Research Associateship.

again based on seemingly reasonable assumptions. In this note we present a result which could be seen as offering another serious criticism of this principle.

Expressed in terms of the predicate calculus Johnson's and Carnap's programmes can be seen as applying simply to a unary language, that is a language with only unary predicates and a countable collection of constants which are intended to exhaust the universe. Whilst Kemeny in 1963 (see [3]) noted that the next step was to generalize the work of Johnson, Carnap et al to higher arity languages, progress along these lines (with the lone exception of [4]) had to wait until recent work of Nix and Paris ([5], [6]). What we shall show is that within the context of *binary* languages a natural generalization of Johnson's Sufficientness Principle is inconsistent with any nontrivial solutions, at least in the presence of another requirement, which is generally assumed. This then provides a further, to our minds strong, criticism of this general principle.

2 Framework

Let L be a language for the predicate calculus with r_i-ary predicates P_i for $i \in \{1, 2, \dots m\}$, constants $a_1, a_2, \dots$ and no function symbols. The intention here is that the constants name all the individuals in the universe, or putting it another way, that we are limiting ourselves to structures in which this is true.

Let FL denote the set of formulae of L, SL the set of sentences of L (with connectives $\neg, \wedge, \vee, \rightarrow, \leftrightarrow$). The 'problem of induction' which we are interested in is the following:

Given some finite consistent subset Γ of SL, and nothing more, what belief should we give to some other θ in SL?

We identify 'belief' with probability, that is, a function w that assigns to each formula from SL a number between 0 and 1, satisfying the following: for any $\theta, \phi \in SL$ and $\psi(x) \in FL$,

(P1) if $\models \theta$ then $w(\theta) = 1$
(P2) if $\models \neg(\theta \wedge \phi)$ then $w(\theta \vee \phi) = w(\theta) + w(\phi)$
(P3) $w(\exists x \psi(x)) = lim_{n \to \infty} w(\bigvee_{i=1}^{n} \psi(a_i))$

We remark that by a theorem of Gaifman [7], w is uniquely determined by its restriction to the quantifier free sentences of L, and any function w on the quantifier free sentences of L satisfying (P1) and (P2) extends to SL so as to satisfy (P3). Hence we can essentially restrict our attention to probability functions w defined on the quantifier free sentences of L.

We also assume that the above problem essentially is the problem of choosing a w on the basis of zero knowledge, i.e. empty Γ, by identifying 'the belief of θ given Γ' with 'the conditional probability of θ given $\bigwedge \Gamma$ according to w', i.e. with

$$\frac{w(\theta \wedge \bigwedge \Gamma)}{w(\bigwedge \Gamma)}$$

(assuming $w(\bigwedge \Gamma) \neq 0$).

Note that whatever w we choose, it will give value 1 to tautologies and value 0 to contradictions. To be able to say anything else, we need some principles which would narrow down the possibilities. The most commonly accepted principle is that of exchangeability, or 'the permutation postulate' in the terminology W.E.Johnson (1924) [8].

The Constant Exchangeability Principle (Ex) If $\theta \in SL$ and θ' is the result of replacing the distinct constant symbols $a_{i_1}, a_{i_2}, \ldots, a_{i_n}$ which occur in θ respectively by distinct $a_{k_1}, a_{k_2}, \ldots, a_{k_n}$ then $w(\theta) = w(\theta')$.

In what follows we shall assume Constant Exchangeability.

3 Unary Case

For this section let us assume that L is purely unary, as indeed was the case in the work of Johnson, Carnap et al. In this case, the Constant Exchangeability principle allows a very pleasing characterization of the w which satisfy it, due to de Finetti (see [9]). Before stating it, we introduce some notation:

Let $P_1, \ldots, P_m$ be the predicate symbols of L, all unary. An atom of L is a formula of the form

$$\bigwedge_{i=1}^{m} P_i^{\epsilon_i}(x)$$

where the $\epsilon_i \in \{0, 1\}$ and $P^1 = P,\ P^0 = \neg P$. The atoms will be denoted $\alpha_h(x),\ h = 1, \ldots, 2^m$. Sentences of the form

$$\bigwedge_{i=1}^{p} \alpha_{h_i}(a_i)$$

where $p \in \mathbb{N}$, are called state descriptions. Note that Gaifman's theorem implies that w is determined by its values on the state descriptions.

Theorem 1 [De Finetti's Representation Theorem] *If w satisfies Ex then there is a probability measure μ on*

$$\mathbb{D}_m = \left\{ \langle x_1, x_2, \ldots, x_{2^m} \rangle \mid x_1, x_2, \ldots, x_{2^m} \geq 0 \ and \ \sum_{i=1}^{2^m} x_i = 1 \right\}$$

such that

$$w\left(\bigwedge_{i=1}^{n} \alpha_{h_i}(a_i) \right) = \int_{\mathbb{D}_m} \prod_{r=1}^{2^m} x_r^{n_r} \, d\mu(\vec{x}),$$

where n_r is the number of elements of $\{i \mid h_i = r\}$ for $r = 1, 2, \ldots, 2^m$.

De Finetti's theorem, although extremely satisfactory (even more so when contrasted with the lack of a similar result for binary predicates to date), does not help us a great deal in choosing a w since there are many of them satisfying exchangeability.

Continuing to consider a language L with finitely many unary predicates $P_1, \dots, P_m$, we come to Johnson Sufficientness Postulate, which does determine a one-parametric family of w, the Carnap Continuum of Inductive Methods, and hence provides some sort of partial answer to our question.

Johnson Sufficientness Postulate JSP $w\left(\alpha(a_{n+1}) \mid \bigwedge_{i=1}^n \alpha_{h_i}(a_i)\right)$ depends only on n and on the number of times that the atom α appears amongst the $\alpha_{h_1}, \dots, \alpha_{h_n}$.

In the unary case which we are considering, JSP implies that all atoms are equivalent in the sense that $w\left(\bigwedge_{i=1}^n \alpha_{h_i}(a_i)\right)$ is a function of the 2^m sizes n_k of the sets $\{i \mid h_i = k\}$ $(k = 1, 2, \dots, 2^m)$ regardless of which atom is associated with which size. This property is referred to as *Atom Exchangeability* (Ax). (JSP has been considered in a weaker form where it is assumed that $w(\alpha(a_{n+1}) \mid \bigwedge_{i=1}^n \alpha_{h_i}(a_i)$ depends on n, on the number of times that the atom α appears amongst the $\alpha_{h_1}, \dots, \alpha_{h_n}$, *and* on α, (cf [10]), in which case Ax does not follow, but a modification of Carnap's contimuum can still be obtained.)

Carnap's continuum of inductive methods was first derived by Johnson in [1] (1932) and then, independently by Kemeny in 1952 (published 1963, see [3]) and Carnap-Stegmuller (1959) (see [11]). It is described by the following theorem.

Theorem 2 *Under the assumptions of Ex and JSP and assuming that the language has at least two predicates, either*

$$w\left(\alpha(a_{n+1}) \mid \bigwedge_{i=1}^{n} \alpha_{h_i}(a_i)\right) = \frac{s + \frac{\lambda}{2^m}}{n + \lambda} \tag{1}$$

for some $0 < \lambda \leq \infty$ *where* s *is the number of times that the atom* α *appears as a conjunct in* $\bigwedge_{i=1}^n \alpha_{h_i}(a_i)$ *or*

$$w\left(\bigwedge_{i=1}^{n} \alpha_{h_i}(a_i)\right) = \begin{cases} 2^{-m} & \textit{if all the } h_i \textit{ are equal} \\ 0 & \textit{otherwise.} \end{cases}$$

Note that (1) uniquely determines w since we have

$$\begin{aligned} &w\left(\bigwedge_{i=1}^n \alpha_{h_i}(a_i)\right) = \\ &\quad = w(\alpha_{h_n}(a_n) \mid \bigwedge_{i=1}^{n-1} \alpha_{h_i}(a_i)) \cdot w(\alpha_{h_{n-1}}(a_{n-1}) \mid \bigwedge_{i=1}^{n-2} \alpha_{h_i}(a_i)) \dots w(\alpha_{h_1}(a_1)). \end{aligned}$$

Carnap's continuum has been widely discussed both by Carnap and later authors (see e.g. [12], [6]). We will only mention one recent development in the debate, due to Nix and Paris [13]. They introduced a principle called the Generalized Principle of Instantial Relevance (GPIR)[1] and showed that just as JSP leads to

[1]The Principle of Instantial Relevance (PIR) says that for atoms $\alpha, \alpha_{h_1}, \dots, \alpha_{h_n}$, $w\left(\alpha(a_{n+2}) \mid \alpha(a_{n+1} \wedge \bigwedge_{i=1}^n \alpha_{h_i}(a_i)\right) \geq w\left(\alpha(a_{n+1}) \mid \bigwedge_{i=1}^n \alpha_{h_i}(a_i)\right)$. PIR follows from Ex. GPIR says that if $\theta(x) \vDash \phi(x)$ then $w(\theta(a_{n+2}) \mid \phi(a_{n+1}) \wedge \psi(a_1, \dots, a_n)) \geq w(\theta(a_{n+2}) \mid \psi(a_1, \dots, a_n))$.

Carnap's continuum, GPIR leads to another 'continuum of inductive methods' defined by

$$w\left(\bigwedge_{i=1}^{n} \alpha_{h_i}(a_i)\right) = 2^{-m} \sum_{r=1}^{2^m} \gamma^{n-n_r}(\gamma+\delta)^{n_r}$$

where n_r is the number of elements of $\{i \mid h_i = r\}$ for $r = 1, 2, \ldots, 2^m$, $\delta + 2^m\gamma = 1$ and $0 \leq \delta \leq 1$.

4 Binary Case

In the rest of this paper we work with a language L containing m predicates $R_1, \ldots, R_m$, all binary. Atoms (denoted $\beta_1(x,y), \ldots, \beta_{2^m}(x,y)$) are now the formulas

$$\bigwedge_{i=1}^{m} R_i^{\epsilon_i}(x,y)$$

and state descriptions (which determine w) are the sentences

$$\bigwedge_{i,j=1}^{p} \beta_{r_{ij}}(a_i, a_j)$$

In [5] and [6] it is argued that apart from Ex, it is reasonable to impose a property analogous to the Atom Exchangeability mentioned above in the unary context. To formulate this property, we need the concept of a *spectrum* associated with a state description:

Given a state description of our binary language

$$\bigwedge_{i,j=1}^{p} \beta_{r_{ij}}(a_i, a_j),$$

let $\mathbf{r}$ be the $p \times p$ matrix with entries r_{ij}. We define the relation I^r to hold between $1 \leq i, j \leq p$ just if $r_{hj} = r_{hi}$ and $r_{ih} = r_{jh}$ for $1 \leq h \leq p$. I^r is clearly an equivalence relation on $\{1, \ldots, p\}$.

The spectrum of $\mathbf{r}$, $S(\mathbf{r})$ is the tuple $\langle | I_1 |, | I_2 |, \ldots, | I_q | \rangle$, where $| I |$ denotes the number of elements of I and $I_1, \ldots, I_q$ are the equivalence classes of I^r arranged in non-increasing order of size.

Spectrum Exchangeability Principle (Sx) If $\mathbf{o}$ and $\mathbf{r}$ from $\{1, 2, \ldots, 2^m\}^{p \times p}$ have the same spectrum then .

$$w\left(\bigwedge_{i,j=1}^{p} \beta_{r_{ij}}(a_i, a_j)\right) = w\left(\bigwedge_{i,j=1}^{p} \beta_{o_{ij}}(a_i, a_j)\right).$$

By analogy with the unary case we formulate

Johnson's Binary Sufficientness Postulate JBSP For a natural number p and $\mathbf{r} \in \{1, 2, \ldots, 2^m\}^{(p+1)\times(p+1)}$,

$$w\left(\bigwedge_{i,j=1}^{p+1} \beta_{r_{ij}}(a_i, a_j) \mid \bigwedge_{i,j=1}^{p} \beta_{r_{ij}}(a_i, a_j)\right) \tag{2}$$

depends only on p and on the number s of k, $1 \le k \le p$ such that

$$r_{p+1i} = r_{ki} \textit{ and } r_{ip+1} = r_{ik} \quad \textit{for all } 1 \le i \le p+1.$$

In the unary case it is easy to see that Johnson's Postulate implies Atom Exchangeability. In the binary case the situation concerning the analogous result (that JBSP implies Sx) is more complicated. Below we shall prove that Sx does in fact follow from JBSP if we also have Ex. Then we shall show that JBSP and Ex hold simultaneously for only two probability distributions w neither of which is an ideal answer to our original question. Since Ex is a widely accepted assumption in this area, the result shows JBSP in a rather unfavourable light.

Given w satisfying JBSP let $g(p, s)$ denote the probabilities defined in (2). For $\mathbf{r} \in \{1, 2, \ldots, 2^m\}^{(p+1)\times(p+1)}$ we have

$$\begin{aligned} w\left(\bigwedge_{i,j=1}^{p+1} \beta_{r_{ij}}(a_i, a_j)\right) = \\ &= w\left(\bigwedge_{i,j=1}^{p+1} \beta_{r_{ij}}(a_i, a_j) \mid \bigwedge_{i,j=1}^{p} \beta_{r_{ij}}(a_i, a_j)\right) \times \\ &\times w\left(\bigwedge_{i,j=1}^{p} \beta_{r_{ij}}(a_i, a_j) \mid \bigwedge_{i,j=1}^{p-1} \beta_{r_{ij}}(a_i, a_j)\right) \times \ldots \\ &\times w\left(\bigwedge_{i,j=1}^{2} \beta_{r_{ij}}(a_i, a_j) \mid \beta_{r_{11}}(a_1, a_1)\right) \times w(\beta_{r_{11}}(a_1, a_1)) \end{aligned}$$

Therefore

$$w\left(\bigwedge_{i,j=1}^{p+1} \beta_{r_{ij}}(a_i, a_j)\right) = g(p, s_p) \times g(p-1, s_{p-1}) \times \ldots \times g(1, s_1) \times g(0, 0)$$

where for $1 \le l \le p$, s_l is the number of k, $1 \le k \le l$ such that

$$r_{l+1i} = r_{ki} \textit{ and } r_{il+1} = r_{ik} \quad \textit{for all } 1 \le i \le l+1.$$

Theorem 3 *JBSP and Ex imply Sx.*

To prove the theorem, we need the following lemma:

Lemma 1 *Let* $\mathbf{r} \in \{1, 2, \ldots, 2^m\}^{p\times p}$ *have spectrum* $\langle\underbrace{1, 1, \ldots, 1}_{p \textit{ times}}\rangle$. *There exists a permutation* σ *of* $\{1, 2, \ldots, p\}$ *such that* $\mathbf{o} \in \{1, 2, \ldots, 2^m\}^{p\times p}$ *defined by*

$$o_{ij} = r_{\sigma^{-1}(i)\sigma^{-1}(j)} \quad \textit{for } i, j \in \{1, \ldots, p\}$$

is such that $\mathbf{o}$ *restricted to* $(k \times k)$ *has spectrum* $\langle\underbrace{1, 1, \ldots, 1}_{k \textit{ times}}\rangle$ *for each* $1 \le k \le p$.

Proof First we shall show that there must exist some $k \leq p$ such that swapping the kth and pth row and the kth and pth column produces $\mathbf{u} \in \{1, 2, \ldots, 2^m\}^{p \times p}$ such that $\mathbf{u}$ restricted to $(p-1) \times (p-1)$ has spectrum $\langle \underbrace{1, 1, \ldots, 1}_{p-1 \ times} \rangle$ If there was no such k then each k could be associated with a pair m_k, q_k such that

$$\langle r_{m_k i}, r_{i m_k} \rangle = \langle r_{q_k i}, r_{i q_k} \rangle \quad for\ i \in \{1, \ldots, p\} - \{k\}$$

and

$$\langle r_{m_k k}, r_{k m_k} \rangle \neq \langle r_{q_k k}, r_{k q_k} \rangle.$$

This produces a graph on $\{1, \ldots, p\}$ with p edges, one for each k, connecting m_k and q_k. Such a graph must contain a cycle, but that is plainly absurd - a vertex m corresponds to the vector $\langle \langle r_{mi}, r_{im} \rangle,\ i = 1, \ldots, p \rangle$ and vertices connected by an edge in this graph correspond to vectors differing precisely on one $\langle r_{mk}, r_{km} \rangle$, with each k used only once.

The lemma now follows since we can generate σ by swapping p with k as above, producing $\mathbf{u}$, then $p-1$ with some k found analogously from $\mathbf{u}$ restricted to $(p-1) \times (p-1)$ etc.

Proof of Theorem 3 Let $\mathbf{r} \in \{1, 2, \ldots, 2^m\}^{p \times p}$ have spectrum $\vec{p} = \langle p_1, \ldots, p_q \rangle$. Let σ be a permutation of $\{1, \ldots, p\}$ such that $\{\sigma^{-1}(1), \ldots, \sigma^{-1}(q)\}$ contains exactly one element of each equivalence class of $I^{\mathbf{r}}$, $\{\sigma^{-1}(q + \sum_{j=1}^{i-1} p_j + 1), \ldots, \sigma^{-1}(q + \sum_{j=1}^{i-1} p_j + p_i - 1)\}$ (which is empty if $p_i = 1$) contains all the remaining elements of I_i and moreover if $\mathbf{o} \in \{1, 2, \ldots, 2^m\}^{p \times p}$ is defined by

$$o_{ij} = r_{\sigma^{-1}(i)\sigma^{-1}(j)} \quad for\ i, j \in \{1, \ldots, p\}$$

then $\mathbf{o}$ restricted to $(k \times k)$ has spectrum $\langle \underbrace{1, 1, \ldots, 1}_{k\ times} \rangle$ for each $1 \leq k \leq q$. That such a σ exists is clear from the lemma. By Ex, we have

$$w\left(\bigwedge_{i,j=1}^{p} \beta_{r_{ij}}(a_i, a_j)\right) = w\left(\bigwedge_{i,j=1}^{p} \beta_{o_{ij}}(a_i, a_j)\right)$$

and by JBSP we have

$$\begin{aligned} w\left(\bigwedge_{i,j=1}^{p} \beta_{o_{ij}}(a_i, a_j)\right) &= g(0,0) \times g(1,0) \times \ldots g(q-1,0) \\ & \quad \times g(q,1) \times \ldots \times g(q+p_1-1, p_1-1) \\ & \quad \times g(q+p_1,1) \times \ldots \times g(q+p_1+p_2-1, p_2-1) \ldots \end{aligned}$$

This shows that $w\left(\bigwedge_{i,j=1}^{p} \beta_{r_{ij}}(a_i, a_j)\right)$ depends only on the spectrum $\vec{p}$, as required.

Theorem 4 *If w satisfies JBSP and Ex then w is either the independent distribution such that for each p and $\mathbf{r} \in \{1, 2, \ldots, 2^m\}^{p \times p}$,*

$$w\left(\bigwedge_{i,j=1}^{p} \beta_{r_{ij}}(a_i, a_j)\right) = (2^{-m})^{p^2}$$

or the distribution defined by

$$w\left(\bigwedge_{i,j=1}^{p} \beta_{r_{ij}}(a_i, a_j)\right) = \begin{cases} 2^{-m} & \textit{if all the } r_{ij} \textit{ are equal} \\ 0 & \textit{otherwise.} \end{cases}$$

Proof By virtue of Theorem 3 our assumptions imply Sx. Hence for all $\mathbf{r} \in \{1, 2, \ldots, 2^m\}^{p \times p}$ with a given spectrum $\vec{p}$, $w\left(\bigwedge_{i,j=1}^{p} \beta_{r_{ij}}(a_i, a_j)\right)$ is the same. For simplicity, we write $w(\vec{p})$ to denote it.

Let $p \geq 3$ and let $\vec{p} = \langle p_1, \ldots, p_q \rangle$ be such that $\sum_{i=1}^{q} p_i = p$ and $1 < q < p$, that is, the spectrum is not $\langle 1, \ldots, 1 \rangle$ nor $\langle p \rangle$. Let $\mathbf{r} \in \{1, 2, \ldots, 2^m\}^{p \times p}$ have spectrum $\vec{p}$ and moreover satisfy the following:

- the equivalence class I_q (by convention, the smallest class) is $\{p - p_q + 1, p - p_q + 2, \ldots, p\}$
- $\mathbf{r}$ restricted to $(p - p_q) \times (p - p_q)$ has spectrum $\langle p_1, \ldots, p_{q-1} \rangle$, i.e. the other equivalence classes are differentiated already within $\mathbf{r}$ restricted to $(p - p_q) \times (p - p_q)$
- $r_{11} \neq r_{p1}$

We have

$$w(\vec{p}) = w\left(\bigwedge_{i,j=1}^{p} \beta_{r_{ij}}(a_i, a_j)\right)$$

and

$$w\left(\bigwedge_{i,j=1}^{p} \beta_{r_{ij}}(a_i, a_j)\right) = g(p-1, s_{p-1}) \times \ \ldots \ \times g(1, s_1) \times g(0, 0)$$

where for $1 \leq l \leq p - 1$, s_l is the number of k, $1 \leq k \leq l$ such that

$$r_{l+1i} = r_{ki} \textit{ and } r_{il+1} = r_{ik} \ \ \textit{for all } 1 \leq i \leq l+1.$$

Note that $s_{p-p_q} = 0$, $s_{p-p_q+1} = 1$, $s_{p-1} = p_q - 1$ Modifying $\mathbf{r}$ by replacing all of the r_{1j}, for $p - p_q + 1 \leq j \leq p$ (note that they are equal) by another element of $\{1, 2, \ldots, 2^m\}$ (so that they remain equal but it is no longer the case that $r_{1p} = r_{ip}$ for $i \in I_1$, $i \neq 1$ produces $\mathbf{o} \in \{1, 2, \ldots, 2^m\}^{p \times p}$ with spectrum

$\vec{o} = \langle p_1 - 1, \ldots, p_q, 1\rangle$ (if $p_1 = p_2 = \ldots p_v$ then $p_1 - 1$ should be moved after p_v to observe the convention of listing spectrum in non-increasing order). We have

$$w(\vec{o}) = w\left(\bigwedge_{i,j=1}^{p} \beta_{o_{ij}}(a_i, a_j)\right)$$
$$= g(p-1, s_{p-1}) \times \ \ldots \ \times g(1, s_1) \times g(0,0)$$

for the same $s_{p-1}, \ldots, s_1$. (This is the case since $r_{ij} = o_{ij}$ for $1 \leq i, j \leq p - p_q$ so s_l remains the same for $1 \leq l \leq p - p_q - 1$ and both for $\mathbf{r}$ and $\mathbf{o}$, $s_{p-p_q} = 0$, $s_{p-p_q+1} = 1, \ldots, s_{p-1} = p_q - 1$. Hence $w(\vec{p}) = w(\vec{o})$ and repeating the same procedure until no equivalence class has more than 1 element yields

$$w(\vec{p}) = w(\langle \underbrace{1, 1, \ldots, 1}_{p\ times} \rangle) \quad \textit{for all } \vec{p} \neq \langle p \rangle.$$

Note that this has been proved for any $p \geq 3$ and $\vec{p} \neq \langle p \rangle$. For $p = 2$, we have only spectra $\langle 1, 1\rangle$ and $\langle 2 \rangle$, and for $p = 1$ only $\langle 1 \rangle$.

For $p \geq 1$ define

$$w_p = (\langle \underbrace{1, 1, \ldots, 1}_{p\ times} \rangle) \quad \textit{and} \quad h_p = w(\langle p \rangle)$$

Clearly, $w_1 = h_1 = 2^{-m}$. For $p \geq 2$, considering in turn some $\mathbf{r} \in \{1, 2, \ldots, 2^m\}^{p \times p}$ with spectrum $\langle p \rangle$, and $\mathbf{r} \in \{1, 2, \ldots, 2^m\}^{p \times p}$ with another spectrum, extended to $\mathbf{r}^+ \in \{1, 2, \ldots, 2^m\}^{(p+1)\times(p+1)}$ respectively so that $p + 1$ does not fall to any equivalence class of $\mathbf{r}$, we see that

$$h_p g(p, 0) = w_{p+1}, \quad w_p g(p, 0) = w_{p+1}.$$

Hence if $w_{p+1} \neq 0$, $g(p, 0) \neq 0$ and $h_p = w_p$.

However, for $p \geq 2$, picking some $\mathbf{r} \in \{1, 2, \ldots, 2^m\}^{p \times p}$ with spectrum other than $\langle p \rangle$, we note that any of its $(2^m)^{2p+1}$ extensions belonging to $\{1, 2, \ldots, 2^m\}^{(p+1)\times(p+1)}$ has spectrum other than $\langle p + 1\rangle$, so

$$w_p = (2^m)^{2p+1} w_{p+1}$$

and consequently only two possibilities arise:

- $w_p \neq 0$ for all $p \geq 2$ and hence $w_p = h_p$ for all $p \geq 1$. This corresponds to the independent distribution described in the proposition.
- $w_p = 0$ for all $p \geq 2$. This corresponds to the other probability function w described in the proposition.

The theorem follows.

We remark that neither solution appears particularly suitable as an answer to our question, since using the independent one means that past information has no effect, and using the other one means that probability 1 is given to all future experiments having the same outcome as the first.

References

[1] Johnson, W.E. (1932) Probability: the deductive and inductive problems, *Mind* Vol.49, pp 409-423

[2] Paris, J.B. (1994) *The Uncertain Reasoner's Companion*, Cambridge University Press

[3] Kemeny, J.G. (1963) Carnap's theory of probability and induction, in *The Philosophy of Rudolf Carnap*, ed. P.A.Schilpp, La Salle, Illinois, Open Court, pp 711-738.

[4] Krauss, P.H. (1969) Representation of Symmetric Probability Models, *The Journal of Symbolic Logic* Vol.34, No.2 pp 183-193

[5] Nix, C.J. and Paris, J.B. (2006) A Note on Binary Inductive Logic, submitted to the *Journal of Philosophical Logic* [`http://www.maths.man.ac.uk/~jeff/#publications`]

[6] Nix, C.J. (2005) *Probabilistic Induction in the Predicate Calculus*, Ph.D. Thesis, The University of Manchester (UK) [`http://www.maths.man.ac.uk/~jeff/#students`]

[7] Gaifman, H (1964) Concerning measures on first order calculi, *Israel Journal of Mathematics*, Vol 2, pp 1-18

[8] Johnson, W.E. (1924) *Logic, Part III. The Logical Foundation of Science*, Cambridge University Press [Reprinted 1964, Dover, New York]

[9] Feller, W. (1966) *An Introduction to Probability Theory and its applications*, Vol II, John Viley & Sons, New York

[10] Zabell, S.L. (1982) W.E.Johnson's 'Sufficientness' Postulate, *The Annals of Statistics*, Vol.10, No.4, pp 1090-1099

[11] Carnap, R. and Stegmuller, W (1959) *Induktiv Logik und Wahrscheinlichkeit*, Springer Verlag, Vienna

[12] Carnap, R. (1980) A basic system of inductive Logic, *Studies in Inductive Logic and Probability*, Vol. II, ed R.C.Jeffrey, University of California Press, 7-155

[13] Nix, C.J. and Paris, J.B. (2005) A continuum of inductive methods arising from a generalised principle of instantial relevance *Journal of Philosophical Logic*, Online First Issue, DOI:10,1007/s 10992-005-9003x, ISSN 0022-3611 (Paper) 1573-0433

[14] Carnap,R. and Jeffrey, R.C. eds. (1971) *Studies in Inductive Logic and Probability*, University of California Press

[15] Carnap, R. (1952) *The Continuum of Inductive Methods*, University of Chicago Press