Reichenbach's Axiom – For the record

J.B.Paris & A.Vencovská
School of Mathematics
The University of Manchester
Manchester M13 9PL

jeff.paris@manchester.ac.uk, alena.vencovska@manchester.ac.uk

July 8, 2014

Abstract

The main purpose of this note is to make accessible a prove of an unpublished result by Haim Gaifman that, assuming Regularity, a probability function satisfies Reichenbach's Principle just and only if every point in $\mathbb{D}_{2^q}$ is a support point of its de Finetti prior.

The following principle has been attributed to Hans Reichenbach after a suggestion by Hilary Putnam, see [2, p120]:

Reichenbach's Axiom, RA

Let $\alpha_{h_i}(x)$ for $i = 1, 2, 3, ...$ be an infinite sequence of atoms of L.[1] *Then for $\alpha_j(x)$ an atom of L,*

$$\lim_{n\to\infty} \left(w \left(\alpha_j(a_{n+1}) \,|\, \bigwedge_{i=1}^{n} \alpha_{h_i}(a_i) \right) - \frac{u_j(n)}{n} \right) = 0 \tag{1}$$

where $u_j(n) = |\{i \,|\, 1 \leq i \leq n \text{ and } h_i = j\}|$.

[1] For an explanation of the notation etc. used in this paper see, for example, [3]. Note however that $\mathbb{D}_q$ in that paper is denoted $\mathbb{D}_{2^q}$ in this current paper.

Informally then this principle asserts that as information of the atoms satisfied by the $a_1, a_2, \ldots, a_n, \ldots$ grows so w should treat this information like a *statistical sample* giving a value to the probability that the next, $n+1$'st, case revealed will be $\alpha_j(a_{n+1})$ which gets arbitrarily close to the frequency of past instances of $\alpha_j(a_i)$.

The following theorem which was stated by Gaifman in [1] though its proof, which it was said would be given in [2], never appeared (due to the inordinate time lag between these two volumes and Gaifman developing new interests in the meantime).

Theorem 1 *Let w satisfy Reg. Then w satisfies RA if and only if every point in $\mathbb{D}_{2^q}$ is a support point of the de Finetti prior μ of w.*

Proof First assume that every point in $\mathbb{D}_{2^q}$ is a support point of μ. By de Finetti's Theorem it is enough to show that if n is large and $m_1, m_2, \ldots, m_{2^q} \in \mathbb{N}$ with sum n then

$$\frac{\int_{\mathbb{D}_{2^q}} (x_j - m_j/n) \prod_{i=1}^{2^q} x_i^{m_i}\, d\mu}{\int_{\mathbb{D}_{2^q}} \prod_{i=1}^{2^q} x_i^{m_i}\, d\mu} \tag{2}$$

is close to zero. We first need to introduce some notation and derive a number of estimates.

For small $\delta > 0$ set

$$E_\delta = \{\vec{x} \in \mathbb{D}_{2^q} \,|\, x_i \geq \delta,\ i = 1, 2, \ldots, 2^q\},$$
$$E_\delta(\vec{c}) = \{\vec{x} \in N_{\delta/2}(\vec{c}) \,|\, \exists\, \vec{y} \in E_\delta\, \exists \lambda \in [0,1], \vec{x} = \lambda\vec{c} + (1-\lambda)\vec{y}\}.$$

Notice that for every point $\vec{d} \in E_\delta(\vec{c})$ and $i = 1, 2, \ldots, 2^q$, if $c_i < \delta$ then $c_i \leq d_i$. Also there is a fixed $\xi > 0$ such that for each $\vec{c} \in \mathbb{D}_{2^q}$ there is a $\vec{d} \in E_\delta(\vec{c})$ such that $E_\delta(\vec{c})$ contains the neighbourhood $N_\xi(\vec{d})$.[2] Hence there is some $\zeta > 0$ such that

$$\forall \vec{c} \in \mathbb{D}_{2^q}\ \mu(E_\delta(\vec{c})) \geq \zeta, \tag{3}$$

[2]It can be checked that a suitable choice, for δ small, is $\xi = 2^{-q-3}\delta$ when $\vec{d}$ is given by $d_i = c_i - (2^q - 1)2^{-q-2}\delta$ for some i for which $c_i = \max\{c_j \,|\, j = 1, 2, \ldots, 2^q\}$ and $d_j = c_j + 2^{-q-2}\delta$ for the remaining $2^q - 1$ coordinates.

since if not we could find a sequence of points $\vec{c}^k \in \mathbb{D}_{2^q}$ with limit point $\vec{c}$ such that $\mu(N_\xi(\vec{c}^k)) \to 0$ whilst, by the assumption on the support points of μ, $\mu(N_{\xi/2}(\vec{c})) > 0$ with $N_{\xi/2}(\vec{c})) \subseteq N_\xi(\vec{c}^k))$ for k large enough.

For $\vec{d} \in E_\delta(\vec{c})$ we have that

$$\begin{aligned} \sum_{i=1}^{2^q} (c_i \log(c_i) - c_i \log(d_i)) &= \\ -\sum_{c_i \geq \delta} c_i \log(1 + (d_i - c_i)c_i^{-1}) + \sum_{c_i < \delta} c_i \log(c_i) - c_i \log(d_i) &\leq 2^{q+1}\sqrt{\delta} \end{aligned} \tag{4}$$

since if $c_i < \delta$ then $c_i \leq d_i$ and $c_i \log(c_i) - c_i \log(d_i) \leq 0$ whilst for $\delta \leq c_i$ in view of $|d_i - c_i| < \delta/2$

$$-c_i \log(1 + (d_i - c_i)c_i^{-1}) \leq c_i \log(2),$$

which is less or equal to $\sqrt{\delta}\log(2) \leq 2\sqrt{\delta}$ in the case of $c_i < \sqrt{\delta}$, and when $c_i \geq \sqrt{\delta}$

$$-c_i \log(1 + (d_i - c_i)c_i^{-1}) \leq -c_i \log(1 - \sqrt{\delta}/2) \leq \frac{c_i\sqrt{\delta}/2}{1 - \sqrt{\delta}/2} \leq 2\sqrt{\delta}.$$

From (4) we now have that for $\vec{d} \in E_\delta(\vec{c})$,

$$\prod_{i=1}^{2^q} d_i^{c_i} \geq e^{-2^{q+1}\sqrt{\delta}} \prod_{i=1}^{2^q} c_i^{c_i}. \tag{5}$$

We now claim that for small $\epsilon > 0$ there exists $\tau > 0$ such that whenever $\vec{c}, \vec{d} \in \mathbb{D}_{2^q}$ and $|\vec{d} - \vec{c}| \geq \epsilon$ then

$$\sum_{i=1}^{2^q} (c_i \log(c_i) - c_i \log(d_i)) \geq \tau.$$

For if not, then since $\sum_i c_i \log(x_i)$ takes its strict maximum on $\mathbb{D}_{2^q}$ at $\vec{x} = \vec{c}$, there would be $\vec{c}, \vec{d}, \vec{c}^k, \vec{d}^k \in \mathbb{D}_{2^q}$ such that $|\vec{d}^k - \vec{c}^k| \geq \epsilon$ for each k, $\vec{c}^k \to \vec{c}$, $\vec{d}^k \to \vec{d}$ but

$$\sum_{i=1}^{2^q} (c_i^k \log(c_i^k) - c_i^k \log(d_i^k)) \searrow 0.$$

In this case $|\vec{d}-\vec{c}| \geq \epsilon$ but

$$\sum_{i=1}^{2^q} c_i \log(c_i) = \sum_{i=1}^{2^q} c_i \log(d_i),$$

contradiction. It follows that the required τ exists and we can conclude that

$$\prod_{i=1}^{2^q} d_i^{c_i} \leq e^{-\tau} \prod_{i=1}^{2^q} c_i^{c_i} \tag{6}$$

whenever $\vec{c}, \vec{d} \in \mathbb{D}_{2^q}$, $|\vec{d}-\vec{c}| \geq \epsilon$.

We now return to the proof that (2) is close to zero. Given small $\epsilon > 0$ let $\tau > 0$ be as in (6). Now pick small $\delta > 0$ such that

$$2^{q+1}\sqrt{\delta} < \tau, \epsilon. \tag{7}$$

Then putting $c_j = m_j/n$ for $j = 1, 2, \ldots, 2^q$,

$$\begin{aligned}
\frac{\int_{\mathbb{D}_{2^q}} (x_j - c_j) \prod_{i=1}^{2^q} x_i^{m_i} \, d\mu}{\int_{\mathbb{D}_{2^q}} \prod_{i=1}^{2^q} x_i^{m_i} \, d\mu} &= \\
&= \frac{\int_{N_\epsilon(\vec{c})} (x_j - c_j) \prod_{i=1}^{2^q} x_i^{m_i} \, d\mu}{\int_{N_\epsilon(\vec{c})} \prod_{i=1}^{2^q} x_i^{m_i} \, d\mu \;+\; \int_{\neg N_\epsilon(\vec{c})} \prod_{i=1}^{2^q} x_i^{m_i} \, d\mu} && (8) \\
&\quad + \frac{\int_{\neg N_\epsilon(\vec{c})} (x_j - c_j) \prod_{i=1}^{2^q} x_i^{m_i} \, d\mu}{\int_{N_\epsilon(\vec{c})} \prod_{i=1}^{2^q} x_i^{m_i} \, d\mu \;+\; \int_{\neg N_\epsilon(\vec{c})} \prod_{i=1}^{2^q} x_i^{m_i} \, d\mu} && (9)
\end{aligned}$$

Concerning (8) we have that

$$\begin{aligned}
\frac{\int_{\neg N_\epsilon(\vec{c})} \prod_{i=1}^{2^q} x_i^{m_i} \, d\mu}{\int_{N_\epsilon(\vec{c})} \prod_{i=1}^{2^q} x_i^{m_i} \, d\mu} &\leq \frac{\int_{\neg N_\epsilon(\vec{c})} \prod_{i=1}^{2^q} x_i^{m_i} \, d\mu}{\int_{E_\delta(\vec{c})} \prod_{i=1}^{2^q} x_i^{m_i} \, d\mu} \\
&\leq \frac{\int_{\neg N_\epsilon(\vec{c})} e^{-n\tau} \prod_{i=1}^{2^q} c_i^{nc_i} \, d\mu}{\int_{E_\delta(\vec{c})} e^{-n2^{q+1}\sqrt{\delta}} \prod_{i=1}^{2^q} c_i^{nc_i} \, d\mu} \quad \text{by (5), (6),} \\
&\leq \frac{e^{-n\tau}}{\zeta e^{-n2^{q+1}\sqrt{\delta}}}. && (10)
\end{aligned}$$

which by (7) is small for large n. Hence (8) is close to

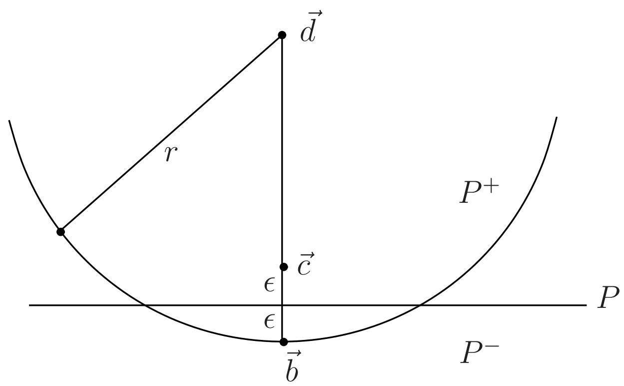
$$\frac{\int_{N_\epsilon(\vec{c})} (x_j - c_j) \prod_{i=1}^{2^q} x_i^{m_i} \, d\mu}{\int_{N_\epsilon(\vec{c})} \prod_{i=1}^{2^q} x_i^{m_i} \, d\mu}$$

which is between $-\epsilon$ and ϵ since $|x_j - c_j| < \epsilon$ over $N_\epsilon(\vec{c})$.

Clearly the inequalities already given in (10) also show that (9) is small for large n and the required result follows.

Turning to the other direction of the theorem suppose that w satisfies Reg and not every point of $\mathbb{D}_{2^q}$ is a support point of the de Finetti prior μ of w. We shall sketch a proof that in this case RA fails in general, even when the sequence $u_j(n)/n$ converges.

Since the set of non-support points of μ form an open set and w satisfies Reg we can find points $\vec{b}, \vec{d} \in \mathbb{D}_{2^q}$ with no zero coordinates with $\vec{b}$ a support point of μ and $\vec{d}$ a non-support point. By considering points on the line joining $\vec{b}, \vec{d}$ we may assume that $\vec{b}$ is close to $\vec{d}$ and, by considering a nearest support point to $\vec{d}$ and then moving a distance in its direction if necessary, that no support point is as close (or closer) to $\vec{d}$ than $\vec{b}$. Let $r = |\vec{b} - \vec{d}| < s/2$ where $s = \min\{b_i, d_i \,|\, i = 1, 2, \ldots, 2^q\}$ and let ϵ be small. In the diagram below let $\vec{c}$ be on the line joining $\vec{b}, \vec{d}$ distance 2ϵ from $\vec{b}$, let the plane P be normal to this line distance ϵ from $\vec{b}$ and let P^+ be the region on the same side of P as $\vec{c}$, P^- its complement. Note that $2r < s \leq c_i$ for each i.



Then

$$\sum_{i=1}^{2^q}(c_i\log(c_i)-c_i\log(b_i)) = \sum_{i=1}^{2^q} -c_i\log\left(1+\frac{(b_i-c_i)}{c_i}\right)$$
$$= \sum_{i=1}^{2^q} -(b_i-c_i)+\frac{(b_i-c_i)^2}{2c_i}+O(\epsilon^3)$$
$$= \sum_{i=1}^{2^q} \frac{(b_i-c_i)^2}{2c_i}+O(\epsilon^3)\le 3s^{-1}\epsilon^2 \qquad (11)$$

since $\sum_{i=1}^{2^q} b_i = \sum_{i=1}^{2^q} c_i = 1$.

On the other hand let $\vec{x}\in P^+$ with $|\vec{x}-\vec{d}|\ge r$ and suppose for the moment that $|x_i-c_i|<c_i$ for each i. The distance from $\vec{x}$ to $\vec{c}$ must be at least $\sqrt{2r\epsilon}$ so

$$\sum_{i=1}^{2^q}(c_i\log(c_i)-c_i\log(x_i)) = \sum_{i=1}^{2^q} -c_i\log\left(1+\frac{(x_i-c_i)}{c_i}\right)$$
$$= \sum_{i=1}^{2^q} -(x_i-c_i)+\frac{(x_i-c_i)^2}{2c_i}-\frac{(x_i-c_i)^3}{3c_i^2}+\ldots$$
$$= \sum_{i=1}^{2^q} \frac{(x_i-c_i)^2}{2c_i}-\frac{(x_i-c_i)^3}{3c_i^2}+\ldots$$
$$\ge \sum_{i=1}^{2^q} \frac{(x_i-c_i)^2}{8c_i} \ \ge\ 2^{-2}rs^{-1}\epsilon. \qquad (12)$$

Furthermore the inequality (12) also holds for any $\vec{x}\in P^+$ with $|\vec{x}-\vec{d}|\ge r$ since the function $\sum_{i=1}^{2^q}(c_i\log(c_i)-c_i\log(x_i))$ is increasing along straight lines emanating from $\vec{c}$. As in the first half of this proof, using (11), (12) it now follows that if

$$\left\langle \frac{u_1(n)}{n},\frac{u_2(n)}{n},\ldots,\frac{u_{2^q}(n)}{n}\right\rangle \ \to\ \vec{c}$$

as $n\to\infty$ then

$$\frac{\int_{\mathbb{D}_{2^q}} x_j\prod_{i=1}^{2^q} x_i^{u_i(n)}\,d\mu}{\int_{\mathbb{D}_{2^q}}\prod_{i=1}^{2^q} x_i^{u_i(n)}\,d\mu}-\frac{\int_{P^-} x_j\prod_{i=1}^{2^q} x_i^{u_i(n)}\,d\mu}{\int_{P^-}\prod_{i=1}^{2^q} x_i^{u_i(n)}\,d\mu} \qquad (13)$$

tends to 0 as $n \to \infty$. But even assuming the limit e_j of the left hand side of (13) exists for $j = 1, 2, \ldots, 2^q$ then, because of the equality with the right hand side of (13), would have to have $\vec{e} \in P^-$ so $\vec{e} \neq \vec{c}$. Either way RA fails, as required. $\blacksquare$

In fact the forward direction of the above proof has shown an ostensibly stronger result, that under the given assumptions RA holds *uniformly*. Precisely:

Corollary 2 *Let w satisfy Reg and suppose that every point in $\mathbb{D}_{2^q}$ is a support point of the de Finetti Prior μ of w. Then for $\epsilon > 0$ there is $k \in \mathbb{N}$ such that for any sequence $\alpha_{h_i}(x)$ of atoms of L and $n \geq k$*

$$\left| w\left(\alpha_j(a_{n+1}) \,|\, \bigwedge_{i=1}^{n} \alpha_{h_i}(a_i) \right) - \frac{u_j(n)}{n} \right| < \epsilon,$$

where $u_j(n) = |\{i \,|\, 1 \leq i \leq n \text{ and } h_i = j\}|$.

We have assumed in Theorem 1 that w satisfies Reg. Clearly this is necessary for RA to make sense *in general* since without Reg we can have the conditional in (1) undefined. Notwithstanding, in [1] Gaifman states a generalization of Theorem 1 appropriate to this case.

If the probability function w satisfies RA and Reg then it satisfies an analogous version of RA for consistent non-tautological $\theta(a_1) \in QFSL$. Namely

$$\lim_{n \to \infty} \left(w\left(\theta(a_{n+1}) \,|\, \bigwedge_{i=1}^{n} \theta^{\epsilon_i}(a_i) \right) - \frac{u(n)}{n} \right) = 0$$

where $u(n) = \sum_{i=1}^{n} \epsilon_i$. To see this notice that the map

$$\vec{x} \in \mathbb{D}_{2^q} \mapsto w_{\vec{x}}(\theta(a_1))$$

is continuous and onto $[0, 1]$ so for μ the de Finetti prior of w the measure ν on $\mathbb{D}_2$ defined by

$$\nu(A) = \mu\{\vec{x} \,|\, \langle w_{\vec{x}}(\theta(a_1)), 1 - w_{\vec{x}}(\theta(a_1)) \rangle \in A\}$$

has every point in $\mathbb{D}_2$ as a support point and by the IP property of the $w_{\vec{x}}$ we have

$$\begin{aligned}\int_{\mathbb{D}_2} w_{\langle x_1,x_2\rangle}\left(\bigwedge_{i=1}^n R_1^{\epsilon_i}(a_i)\right) d\nu(\langle x_1,x_2\rangle) &= \int_{\mathbb{D}_2} x_1^{n_1}x_2^{n_2}\, d\nu(\langle x_1,x_2\rangle)\\ &= \int_{\mathbb{D}_{2^q}} w_{\vec{x}}\left(\bigwedge_{i=1}^n \theta^{\epsilon_i}(a_i)\right) d\mu(\vec{x})\end{aligned}$$

where $n_1 = \sum_{i=1}^n \epsilon_i$, $n_2 = n - n_1$. The required conclusion now follows by applying Theorem 1 to this ν.

Notice in particular then that in this case (and trivially if θ is a tautology)

$$\lim_{n\to\infty} w\left(\theta(a_{n+1}) \mid \bigwedge_{i=1}^n \theta(a_i)\right) = 1.$$

References

[1] Carnap, R., A Basic System of Inductive Logic, in *Studies in Inductive Logic and Probability, Volume I*, Eds. R.Carnap & R.C.Jeffrey, University of California Press, 1971, pp33-165.

[2] Carnap, R., A Basic System of Inductive Logic, in *Studies in Inductive Logic and Probability, Volume II*, ed. R.C.Jeffrey, University of California Press, 1980, pp7-155.

[3] Paris, J.B., Pure Inductive Logic, in *The Continuum Companion to Philosophical Logic*, Eds. L.Horsten & R.Pettigrew, Continuum International Publishing Group, London, 2011, pp428-449.