

## 因果帰納の二要因ヒューリスティクス・モデル

服部 雅史

In regard to the causal induction based on contingency information, the probabilistic contrast, or the difference between the probability of the effect ( $e$ ) in the presence of the cause ( $c$ ), and the probability of the effect in the absence of the cause have been regarded as the appropriate measure of perceived causality in many studies. This article proposes a new model of causal induction which is called the dual-factor heuristics (DH) model. The DH model is based on the two important factors of causal inference:  $P(e|c)$ , the predictability of the effect, and  $P(c|e)$ , the compatibility of the cause. In an experiment, participants made causal judgments from sequential presentations of information about occurrences and non-occurrences of an effect in the presence and absence of possible causes. Participants' judgments were better predicted by the DH model than models based on the probabilistic contrast. Effects of the probabilistic contrast, however, barely detected, and it was considered to be caused by mixed strategies of the participants. The results were discussed from the viewpoint of adaptive rationality.

Keywords: causation (因果性), contingency (随伴性), the probabilistic contrast model (確率対比モデル), the power PC theory (パワー PC 理論), evolutionary rationality (進化的合理性)

### 1. はじめに

ひとは、世界で起こる出来事を個別の事象としてではなく、因果関係というつながりによって結合したまとまりとして知覚する。火事が起きればその原因を探ろうとするし、煙草を吸いながら肺癌のことを気にかけるかもしれない。近代哲学においては、David Humeにより因果的効力の実在性が否定されて以来、数多くの議論がなされてきた。心理学においても、認知心理学者は共変動 (covariation) 情報と因果 (相関) 関係認知の問題として (e.g., Smedslund, 1963), 動物心理学者は因果 (随伴性) 規則の学習の問題として (e.g., Rescorla, 1968), 社会心理学者は原因帰属の問題として (e.g., Kelley, 1967), 古くからこのテーマに関わってきた。

本研究の目的は、確率的情報に基づいた因果関係の帰納に関するモデルを提案し、その記述的妥当性

を確かめることである。ここでは、因果性に関する哲学的議論に踏み込む余裕はないが、人間の認知の記述的理論を考える場合、それはあまり問題とならない。確かに、社会的に共有された知識としての科学的因果関係を分析的に突き止めることは、人類にとって重要な所業であるが、その一方で、個人レベルのリアルタイムで進行する日常生活においては、できるだけ時間とコストをかけず、だいたいうまくいく結論を出すことは、適応的観点から極めて重要である。有限の記憶容量と処理能力しか持たない人間が、現実的な時間内に有用な結論を出さなければならないという制約の中においては、「だいたい〜である」という確率的情報に基づく推論は必然的要求とも言えよう。

原因の候補となる事象と結果事象の共変動情報からどのように因果関係が帰納されるかについての確率的モデルとして、認知心理学、動物の条件づけ学習などの異なる観点から、いずれも以下の  $\Delta P$  に着目するモデル (随伴性モデル contingency

model) が提案されてきた (Jenkins & Ward, 1965; Rescorla, 1968).

$$\Delta P = P(e|c) - P(e|\bar{c}). \quad (1)$$

ここで  $c$ ,  $e$  はそれぞれ原因 (の候補), 結果を表し,  $c$  の上のバーは否定を表す. すなわち, このモデルは, 原因が起きたときに結果が起きた確率 (例えば, 薬を飲んだときに病気が直った確率) と, 原因が起きなかったときに結果が起きた確率 (薬を飲まなかったときに病気が直った確率) との差を考え, その差が 0 より十分大きい場合には因果関係があると考え, そうでない場合は因果関係がないとするものである.

Cheng & Novick (1990, 1992) は, この随伴性モデルに, 当該事象に関する文脈によって決定されるとされる**焦点集合** (forcal set) の概念を取り込み, **確率対比モデル** (probabilistic contrast model; 以下では PC モデル) を提案した<sup>1)</sup>. このモデルはまた, 「原因」と「可能化条件 (enabling condition)」(以降, 単に「条件」とよぶ) の区別も説明した (Cheng & Novick, 1991). 例えば火事が起きたとき, 「煙草の火」を原因と言うことはあっても, 通常, 「酸素」は原因ではなく単なる「条件」と言われる. しかし, 煙草の火も酸素も他の要因と同時に存在することで十分条件となりうること, また, それが必要条件となる個別事例を考えることができることなど, 論理的ステータスにおいて両者に区別はない. つまり, 原因と条件の区別は, 必要性や十分性という論理的概念では捉えられない問題である (e.g., Hart & Honoré, 1959/1985).

PC モデルによれば, 原因の候補と考えられる事象  $c$  が存在するとき, 当該の焦点集合において  $\Delta P \gg 0$  のとき  $c$  は原因と判断される. また, その現在の焦点集合において  $P(c) = 1$  であり, かつ別の焦点集合において  $\Delta P \gg 0$  のとき,  $c$  は単なる条件と判断される. さらに, 上記のいずれにも当てはまらない場合, すなわち,  $\Delta P \simeq 0$  であるか, または, 当該の焦点集合において  $P(c) = 1$  であっ

ても, 他のいかなる焦点集合においても  $\Delta P \simeq 0$  の場合,  $c$  は因果的無関連要因と判断される.

## 2. 二要因ヒューリスティックス・モデル

上記の PC モデルとは別の観点から因果推論の確率モデルを考えることも可能である. そもそも因果推論の意義は, 1 つには, 将来起こるであろう結果を予測・予期することにより, 将来に備え, 環境により適応したよりよい生活を送ることにあると考えられる. ある事象  $c$  が起きたときに別の事象  $e$  が起きる確率が高い場合,  $c$  は  $e$  の生起の予測に役立つ. そこで, 原因  $c$  から結果  $e$  の**予測可能性** (predictability) は,  $P(e|c)$  で定義されるとする. この観点からは, 予測可能性  $P(e|c)$  が高いほど,  $c$  が  $e$  の原因らしいと考えられる.

一方, 日常的因果関係には, 「原因のないところに結果は起こらない」という暗黙の前提がある. これを確率的表現によって表すと  $P(\bar{c}, e) \simeq 0$  となる. しかし,  $P(\bar{c}, e)$  の変動範囲が  $0 \leq P(\bar{c}, e) \leq P(e)$  であることを考えあわせると, 重要なのは  $P(\bar{c}, e)$  の値自体の小ささより, むしろ  $P(e)$  と比較してどれくらい小さいかという問題であることがわかる. すなわち,  $P(\bar{c}, e)/P(e) = P(\bar{c}|e)$  が小さければ小さいほど, 日常的意味における因果律にあてはまることになる. そこで,  $1 - P(\bar{c}|e) = P(c|e)$  を原因の結果に対する**適合性** (compatibility) と定義し, この値が高いほど, 原因候補事象  $c$  が結果  $e$  の原因として適合的であると考えられる.

以上の議論から, 因果関係の帰納のためには, 原因  $c$  から結果  $e$  の予測可能性  $P(e|c)$ , および原因の結果に対する適合性  $P(c|e)$  という 2 つの要因が強く関係しており, これら 2 つの確率がともに高いときは原因  $c$  と結果  $e$  の因果関係が強く認知され, いずれも低いときは因果関係がないと判断されると予測される. では, 予測可能性と適合性の一方のみが高くもう一方が低い場合はどうなるであろうか.

結果を引き起こすための「条件」とは, 当該の文脈において, ごく普通のありふれた事象である. ありふれているということは, その存在自体が結果を予測する可能性, すなわち条件から結果の予測可能性は低いと言えるだろう. 一方, 「条件」という単語には「必要条件」の響きがあることから, 原因の場合と同様, 「条件が満たされないところに結果は起こらない」とされ, 条件の結果に対する適合性は

1) その後, このモデルはパワー PC 理論 (power PC theory; Cheng, 1997) として発展的展開を見せている. この理論の中で彼女は, 随伴性の指標として  $\Delta P$  よりも  $\Delta P/[1 - P(e|\bar{c})]$  を使うことを提案しているが, この提案を否定する実験結果 (Lober & Shanks, 2000) も報告されている. 本研究では, どちらの指標によってもモデルの予測は変わらない刺激を用いたため, 両者を敢えて区別せず PC モデルという用語を用いる.

表1 二要因ヒューリスティックス (DH) モデルによる因果帰納の結果予測

		$P(c e)$	
		高	低
$P(e c)$	高	原因条件	無関連無関連
	低		

高いのではないかと考えられる。したがって、予測可能性  $P(e|c)$  が低く、適合性  $P(c|e)$  が高い場合に、 $c$  は条件と判断されると予想される。逆に、予測可能性  $P(e|c)$  は高いが適合性  $P(c|e)$  が低い場合は、日常的な意味でのナイーブな因果律を満たしていないことから、因果関係は認知されにくいと考えられる。

以上の予測をする因果帰納の確率モデルを二要因ヒューリスティックス・モデル (dual-factor heuristics model; 以下では DH モデル) と呼ぶ。原因、条件、因果的無関連要因に関する DH モデルの予測をまとめると、表1のようになる。

そこで、DH モデルの記述的妥当性を、従来の PC モデルと比較する実験によって確認することにした。以下では、その実験を紹介する。なお、両モデルの予測に関わる  $\Delta P$ 、 $P(c)$ 、 $P(e|c)$ 、 $P(c|e)$  を確率指標、または単に指標とよぶ。

### 3. 実験方法

#### 3.1 被験者

被験者は、大学生および大学院生計 27 名であった。5 名は個別に実験を行い、残る 22 名はそれぞれ 5 名、8 名、9 名の小グループで行った。

#### 3.2 課題

実験課題は、コンピュータ上に時系列的に提示される原因候補事象と結果事象の生起・不生起の組み合わせを一場面ずつ観察し、最後に、その原因候補事象が原因・条件・無関連要因のいずれであるかを強制選択法により判断するというものであった。被験者にとって「原因」と「条件」の区別が明確になるように、Cheng & Novick (1991) による教示の該当部分をそのまま用いた。その内容は、校庭で遊ぶ少女を少年が背後から押し、少女が転んで膝にけがをしたとき、原因は「彼が押したこと」であり、「地球の重力」や「地面が堅いこと」などは条件と考えられる、というものであった。

各セッションの最初に植物に花を咲かせる作用を

持つ可能性のある 3 種類の薬品 (農薬) が提示された。被験者は、その中から自由に 1 つを選択し、セッションの最後にその薬品についてのみ因果性判断をした。セッションを構成する各試行において、被験者は、自分が選択した薬品を含む 3 つの薬品がそれぞれ与えられる／与えられない場面と、植物の花が咲く／咲かない場面が同時に提示された。被験者は自分のペースでそれを観察した。被験者が選択した薬品は、事象の生起頻度が統制されたが (後述)、選択しなかった 2 つの薬品については、それぞれ生起確率 1/2 でランダムに提示された。セッションの終了時の因果性判断の際、その判断に対する確信度をマウスを使って無段階で回答した。確信度は、0 以上 100 以下の整数に変換された。

各セッションに対応する刺激は、表 2 に示す 13 種類である。原因候補事象の生起、不生起をそれぞれ  $c$ 、 $\bar{c}$  と表し、結果についてはそれぞれ  $e$ 、 $\bar{e}$  と表すと、事象の組み合わせとしては、 $ce$ 、 $c\bar{e}$ 、 $\bar{c}e$ 、 $\bar{c}\bar{e}$  の 4 種類が考えられる。これらの事象頻度により確率指標値が定まり、それによって各モデルの予測が決定する。表 2 の刺激 ID は、アルファベットの 1 文字目が DH モデル、2 文字目が PC モデルの予測をそれぞれ表しており、C、N、I の文字はそれぞれ原因、条件、無関連要因を意味する。例えば CN2 は、DH モデルが「原因」、PC モデルが「条件」という反応を予測する刺激である。アルファベットの後に続く数字 (1 または 2) は、以下で述べる刺激バリエーションの識別子である。

DH モデルに関しては、確率値の高い値は、.9 以上とし、低い値は .5 とした。低い値を 0 に近い値にしなかったのは、例えば  $P(e|c) = .1$  とすると  $P(\bar{e}|c) = .9$  となることから、逆に  $c$  が  $e$  の抑制的原因と判断されてしまうのを回避するためである。PC モデルに関しては、 $\Delta P$  が .5 以上のものを  $\Delta P \gg 0$  を満たすとし、他の制約条件と競合しない限り、1.0 に近い値になるようにした。

DH モデルが原因と予測するのは、 $P(e|c)$  と  $P(c|e)$  の両方が高い場合であるが、「高い」というのは相対的な概念であるので、確率としての最高値の 1.0 の場合以外に、ややノイズを含むバリエーション (指標値がそれぞれ .9 程度のもの) を設けた。CC および CN に関しては、こうして 2 種類のバリエーションが設定された。CI については、刺激の制約上、試行回数を非常に多く (最低 100 試

表2 各刺激の事象数と各種の確率指標, 被験者の反応頻度, 回答に対する確信度

ID	事象数				確率指標				反応頻度			確信度	
	ce	cē	cē	cē	$\Delta P$	$P(c)$	$P(e c)$	$P(c e)$	C	N	I	M	SD
CC1	10	0	0	10	1.00	.50	1.00	1.00	23 <sup>++</sup>	4 <sup>-</sup>	0 <sup>--</sup>	82.3	18.0
CC2	18	2	2	18	.80	.50	.90	.90	12	9	6	63.2	22.4
CN1	20	0	0	0	-	1.00	1.00	1.00	20 <sup>++</sup>	6	1 <sup>--</sup>	73.6	26.4
CN2	9	1	0	0	-	1.00	.90	1.00	14 <sup>+</sup>	10	3 <sup>-</sup>	62.2	14.7
CI1	81	9	9	1	.00	.90	.90	.90	9	6	12	59.2	28.2
NC1	5	5	0	5	.50	.67	.50	1.00	4 <sup>-</sup>	8	15 <sup>+</sup>	57.3	25.2
NC2	5	5	0	10	.50	.50	.50	1.00	5	9	13	45.3	23.4
NN1	5	5	0	0	-	1.00	.50	1.00	5	5	17 <sup>++</sup>	50.4	24.1
NI1	19	19	1	1	.00	.95	.50	.95	1 <sup>--</sup>	4 <sup>-</sup>	22 <sup>++</sup>	59.7	23.7
NI2	9	9	1	1	.00	.90	.50	.90	1 <sup>--</sup>	7	19 <sup>++</sup>	55.2	25.7
IC1	5	0	5	5	.50	.33	1.00	.50	5	2 <sup>--</sup>	20 <sup>++</sup>	59.1	21.7
IC2	3	0	3	12	.80	.17	1.00	.50	8	9	10	55.2	22.0
II1	3	3	3	3	.00	.50	.50	.50	3 <sup>-</sup>	1 <sup>--</sup>	23 <sup>++</sup>	65.2	21.7

注) 反応頻度における ++ および -- は,  $\chi^2$  検定の結果 1%水準で有意差が認められた上で行った残差分析において, 1%水準でそれぞれプラスおよびマイナス方向の有意差が認められたことを示す. 同様に, + および - は,  $\chi^2$  検定で 5% または 1%水準で有意差が認められ, 残差分析で 5%水準でそれぞれプラスおよびマイナス方向の有意差が認められたことを示す.

行) する必要があるため, 被験者の負担を考慮して 1種類のみとした. NCに関しては,  $P(e|c) = .5$  と  $\Delta P \geq .5$  を満たし, かつノイズを含む刺激は存在しないため, 単に事象数の異なる 2種類のバリエーションを用意した. NN は, 各事例数を単純に比例配分的に増減したバリエーションしか存在しないため, 複数用意する意義が小さいと判断し 1種類のみとした. NI のノイズを含まない刺激は,  $c\bar{e}$  の事象数が 0 である必要があるが, それは他の条件と競合して作成不可能であったので, ノイズの多/少の 2種類を用意した. IC については, PC モデルによる  $\Delta P$  を高めるという条件から,  $P(e|c)$  の低いバリエーションは作成できなかったため, 代わりに  $\Delta P$  の異なるものを用意した. IN に相当する刺激は制約を満たす刺激が存在しないため用意できなかった<sup>2)</sup>. II は, NN と同じ理由から 1種類のみとした. 以降, 刺激に 2つ以上のバリエーションがある場合, 最初の刺激を第一刺激, 2番目の刺激を第二刺激と呼ぶ.

### 3.3 手続き

実験は, 表 2 の 13 刺激に加え, 被験者全員が最初に実施するダミー刺激 (事象数は  $ce, c\bar{e}, \bar{c}e, \bar{c}\bar{e}$

2) この場合,  $P(c) = 1, P(e|c) = \text{LOW}, P(c|e) = \text{LOW}$  という 3 つの条件を満たす必要がある. 確率の定義より,  $P(e) = P(c, e) + P(\bar{c}, e)$  であるが,  $P(c) = 1$  の制約から  $P(e) = P(c, e)$  となる. よって  $P(c|e) = P(c, e)/P(e) = 1 \neq \text{LOW}$ . 以上より, これらの制約をすべて満たすことは不可能である.

の順に, 9, 1, 1, 9) を 1 つ用意し, 計 14 セッションから構成された. ダミー刺激を除くセッションの実施順序, およびセッション内の試行の順序は被験者毎にランダム化された. 各事象には, セッション間で全く関連性がないことが教示された.

## 4. 実験結果および考察

### 4.1 刺激別の反応頻度

各刺激に対する被験者の反応頻度は, 表 2 に示されている. 刺激別の頻度の偏りに 1%水準で有意差が見られたのは (以下の括弧内は  $\chi^2[2, N = 27]$  値), CC1 (33.56), CN1 (21.56), NN1 (10.67), NI1 (28.67), NI2 (18.67), IC1 (20.67), II1 (32.89) の 7 刺激, また, 5%水準で有意差が見られたのは, CN2 (6.89), NC1 (6.89) の 2 刺激であった. 残る CC2 (2.00), CI1 (2.00), NC2 (3.56), IC2 (0.22) については有意差が見られなかった. 有意差が見られた 9 刺激について残差分析をしたところ, 表 2 に示すような結果となった. 有意に反応数が多い/少ない反応には ++/-- ( $P < .01$ ) または +/- ( $P < .05$ ) がそれぞれ付けられている.

まず, 両モデルが同じ結果を予測する刺激について見ると, CC1 および II1 は, 両モデルの予測通り, それぞれ C および I の反応が明らかに多くなった ( $P < .01$ ). CC2 については, 最頻反応は予測通りであったが, 有意差は見られなかった. しかし, NN1 では両モデルの予測に反して, 「無関連」が多

かった。

次に、両モデルが異なる結果を予測する刺激について順に見ていく。まず、DHモデルが原因と予測する刺激について見ると、PCモデルが条件と予測するCN1, CN2については、いずれもDHモデルの予測通り原因反応が最も多く、いずれも残差分析において有意差が見られた(順に  $P < .01$ ,  $P < .05$ )。PCモデルが無関連と予測するCI1については有意差は見られなかった。次に、DHモデルが無関連と予測するものについて見ると、IC1では明らかにDHモデルを支持する結果が得られた( $P < .01$ )。しかし、IC2では有意差は認められなかった。最後に、DHモデルが条件と予測する刺激については、NC1, NN1, NI1, NI2では、いずれも無関連反応が多く( $P_s < .01$ , ただしNC1のみ  $P < .05$ )、NC2でも有意差はないがその傾向が見られることがわかる。

以上をまとめると、DHモデルが原因または無関連と予測する刺激については、CI1を除いてその予測通りの反応傾向が見られ、一部(CC2, IC2)を除いて有意差が見られた。一方、DHモデルが条件と予測する刺激については、すべて無関連反応が多く、NC2以外では有意差が見られた。

実験結果を見ると、DHモデルの「条件」という予測については、再考の余地がありそうである。表1の  $P(e|c)$  が低く  $P(c|e)$  が高い場合は、「条件」ではなく「無関連」にモデルを修正するのが妥当かも知れない。また、両モデルの予測が一致する刺激のうち、予測が外れたのは両者が「条件」と予測する場合のみであった。PCモデルが条件と予測する要因は、それが別の何らかの焦点集合において原因と認定されていることが前提となるが、今回の実験では、その点は厳密に統制できなかった。したがって、CNとNNにおけるモデル評価に関しては、その意味で適切ではなかったかも知れない。

#### 4.2 刺激に含まれるノイズの影響

CCおよびICに関しては、第一刺激においてのみDHモデルの予測通りの有意差が見られた( $P < .01$ )。CNについては、いずれにおいてもDHモデルの予測通りの有意差が見られたが、ノイズを含む第二刺激(5%水準)に比べ、第一刺激の方が明らかな差が見られた(1%水準)。一方、NIでは、両者でPCモデルの予測通りの有意差が見

られ( $P < .01$ )、NCについては、第一刺激のみで両者の予測と異なる無関連反応が有意に多かった( $P < .05$ )。ただし、どの刺激に関しても、第一刺激と第二刺激において(有意差の有無は別とすれば)同様の反応傾向が見られた。

CC2において有意差が消失したことから、 $P(e|c)$ と $P(c|e)$ の両方が.9という設定は、ノイズが大きすぎた可能性がある。そうだとすれば、有意差が見られなかったCI1についても、ノイズを少なくすることで、DHモデルを支持する結果が得られる可能性があるだろう。ただし、IC1とIC2はDHモデルの指標値は等しい( $\Delta P$ の値が異なる)ことから、この2者間における結果の違いを説明するには、PCモデルの影響を考えざるを得ない。しかし同時に、 $\Delta P$ が.8にまで達しているにも関わらず、原因反応はチャンスレベルにも満たないことから、単純にPCモデルを支持する結果とも言えない。

#### 4.3 選択頻度の対数線形分析

以上、刺激別に分析を進めてきたが、結果を統合的に捉えるために、DHモデル[D]とPCモデル[C]の各予測カテゴリ、および被験者の因果性判断[J]の3要因に関して対数線形モデルを適用した<sup>3)</sup>。ただし、INの組合せに相当する刺激が存在しないため、モデルのパラメータ推定値を求められない<sup>4)</sup>。そこで、PCモデルが「条件」と予測する刺激(CNとNN)を分析対象から除外した。これは、§4.1で述べた刺激の妥当性に対する疑問を考慮に入れた判断である。また、以下の選択基準に従い、第一刺激または第二刺激の一方を分析対象とした。CCについてはノイズを含まないCC1, CN1を分析対象とし、NIについてはノイズの少ないNI1を採用した。NC, ICについては、ア・プリオリな判断を避け、尤度比適合度の高くなる方を選択した。結果としてNC2, IC2が採用された。以上より、[D]と[J]がそれぞれC, N, Iの3水準、[C]はC, Iの2水準となり、分析対象となった被験者の反応頻度データは、 $3 \times 3 \times 2$ の3次元分割表を構成する。

3) 因果性判断が刺激間で独立と仮定すれば、この実験は独立多項抽出(independent multinomial sampling)とみなせるため、[D]と[C]を含む対数線形モデルによる最尤推定が可能となる(e.g., Agresti, 1990, p.169)。

4) 構造的ゼロ(structural zero)を含む多次元クロス表については、この場合のように、対数線形モデルのパラメータの最尤推定値が求められない場合がある(Bishop, Fienberg, & Holland, 1975)。

ここで、 $[D]$  と  $[C]$  は説明変数であり周辺度数が固定されているため、モデルには  $[D \times C]$  の交互作用項を含める必要がある。要求を満たす階層モデルの中で、飽和モデルを除き有意水準 5% で棄却されなかったのは、次のモデルのみであった、 $G^2(4, N = 162) = 7.39, P = .12$ 。

$$\log m_{ijk} = \mu + \lambda_i^D + \lambda_j^C + \lambda_k^J + \lambda_{ij}^{DC} + \lambda_{ik}^{DJ} + \lambda_{jk}^{CJ} \quad (2)$$

ここで、 $m_{ijk}$  は各セルの期待度数 ( $i, j, k$  はそれぞれ  $[D], [C], [J]$  の水準を表す)、 $\mu$  は  $\log m_{ijk}$  の平均、 $\lambda_x^X$  は変数  $X$  の水準  $x$  における主効果、 $\lambda_{xy}^{XY}$  は変数  $X$  と  $Y$  の各水準  $x, y$  における交互作用を表している。式 (2) のパラメータ  $\lambda_i^D, \lambda_j^C, \lambda_{ij}^{DC}$  には説明変数の度数が、 $\lambda_k^J$  には刺激種を無視した場合の各判断の周辺度数がそれぞれ反映されているだけなので、いずれにもあまり意味はない。残る  $\lambda_{jk}^{DJ}$  と  $\lambda_{jk}^{CJ}$  が、それぞれ DH モデルと PC モデルのいわば影響力 (予測力) を表している。DH モデルまたは PC モデルが「原因」と予測するとき、それぞれの予測通りの反応の頻度を規定するパラメータ  $\lambda_{11}^{DJ}, \lambda_{11}^{CJ}$  の最尤推定値 (MLE) は、順に 1.09, 0.60 となった。このことから、「原因」反応に対する影響力は、DH モデルの方がかなり大きい ( $e^{1.09-0.60} = 1.63$  倍) ことがわかる。また、「無関連」反応に関する MLE は、 $\lambda_{22}^{DJ} = 0.41, \lambda_{22}^{CJ} = 0.77$  となり、これは PC モデルの方が大きかった (1.43 倍)。両モデルの予測が「原因」と「無関連」に分かれる刺激に関して、最尤推定パラメータに基づいて C/I 反応頻度のオッズ比を計算すると、CI1 では 1.02, IC1 では 0.96 となり、いずれにおいても僅かながら DH モデルの予測する反応の方が多くなることが示された。

以上より、総合的には DH モデルの方が予測力が強いことが分かるが、PC モデルの影響も無視できない程度には存在することが示された。

#### 4.4 確信度の分析

因果関係の判断の際の確信度の平均値および標準偏差は、表 2 に示されている。対数線形分析の場合と同じデータ (図 1 参照) を用い、DH モデルと PC モデルの予測を各因子とし、二元配置分散分析を行ったところ、交互作用が有意であった、 $F(2, 52) = 12.4, P < .01$ 。そこで、各水準ごとに

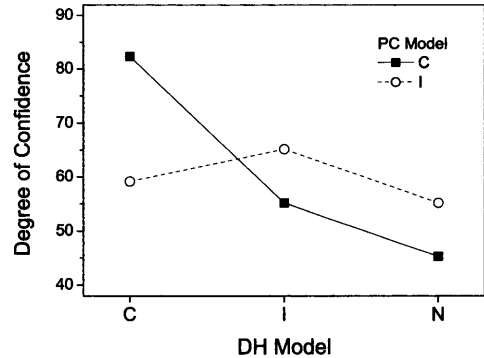


図 1 各刺激に対する判断の確信度評定平均値。横軸は DH モデル、線は PC モデルの予測に従った刺激分類を示す。

単純主効果を分析したところ、PC モデルは DH モデルが「原因」の場合のみ、また、DH モデルは PC モデルが「原因」の場合のみ、それぞれ 1% 水準で有意となった (PC モデル:  $F(1, 78) = 16.56, 3.07, 3.02$ , DH モデル:  $F(2, 104) = 22.62, 1.55$ , いずれも原因, 無関連 [, 条件] の順)。PC モデルが「原因」の水準における多重比較 (Tukey の HSD 法) の結果、DH モデルが「原因」の場合に他の場合よりも高かった、 $MSE = 424.5, P < .01$ 。

以上の結果は、両モデルが原因と予測する場合に、特に確信度が高くなることを示している。両方のモデルに支持されることによって確信度が高まったということは、被験者が因果性判断の際に、随伴性を含む複数の情報を考慮に入れていた可能性を示唆している。つまり、パフォーマンスとしては DH モデルの予測に従ったとしても、その判断過程においては、判断の裏づけとなる他の情報も補助的に利用した可能性がある。

#### 5. 総合的考察

各刺激に対する最頻反応を見る限り、「条件」に関する予測以外、DH モデルの予測はほぼ支持されたと言える。この結果は、本研究において新たに提案された因果性判断のモデルの妥当性を十分に裏づけるものである。一方、対数線形分析および確信度分析の結果は、弱いながらも PC モデルの影響が確かに存在することを示した。以下では、まず、DH モデルに対する肯定的結果の含意について考察し、続いて、PC モデルの影響や「条件」の問題などの DH モデルに対する否定的結果の含意について考察する。

DH モデルは、 $P(e|c)$  と  $P(c|e)$  を指標としている。しかし、 $P(e|c)$  に関して言えば、予測可能性の規範的観点からは、その値自体よりも (PC モデルが主張するような)  $P(e|\bar{c})$  または  $P(e)$  と比較した大小関係を論じない限り、有意味性が保証されない点に注意する必要がある。なぜならば、例えば、仮に  $c$  の起きたときの  $e$  の確率  $P(e|c)$  が .9 という高い値だったとしても、そもそも  $e$  の起こる確率  $P(e)$  (ベースレート) が .9 であれば、 $c$  に関係なく常に  $e$  が高い確率で起るといだけのことであることから、規範的には、 $c$  の原因としての妥当性は低いと考えるべきだからである。同様に、 $P(c|e)$  の大小は  $P(c|\bar{e})$  あるいは  $P(c)$  と比較されるべきだと言える。それなのに敢えてそれを考慮しないというこの種の「手抜き」は、単なるバイアスと捉えられるべきではない。もし、因果叙述文において**稀少性仮説** (rarity assumption; Oaksford & Chater, 1994) が成立しているとすれば、すなわち、実際に因果関係が問題になるような多くの日常的場面においては  $c$  や  $e$  の生起確率が低いとすれば、より手間のかかる  $P(e|c)$  と  $P(e)$  の対比を計算するより、 $P(e|c)$  そのものの値を見る方が、現実的に有効なヒューリスティックス (Gigerenzer, Todd, & The ABC Research Group, 1999) になると考えられる。つまり、 $P(e|c)$  の値を指標値として用いることは、適応的観点から合理的方略と考えることが可能である。

さらに、DH モデルは、 $P(e|c)$  と  $P(c|e)$  の両方が高いときに因果性が知覚されることを主張する。 $P(e|c) = 1$  かつ  $P(c|e) = 1$  のとき  $c$  と  $e$  の関係は  $c \leftrightarrow e$  という実質等値 (双条件文) に一致し、また、因果叙述文がしばしば条件文によって述べられることを考えると、DH モデルの予測は、条件文が双条件的に解釈されるという従来の心理学や言語学における知見に対し、適応的合理性の観点から示唆を与えうるだろう。また、Hattori (in press) は、稀少性と**双条件性** (biconditionality) を仮定することにより、Wason 選択課題における「誤答」の適応的合理性を明らかにしたが、DH モデルはその結果とも整合的である。

次に、 $P(c|e)$  は、§2 で述べた通り、「原因のないところに結果は起こらない」という日常的意味での因果律の反映を意図してモデルに組み込まれたが、観点を変えると、この指標は、結果から逆に原因を

推測する因果推論、すなわち原因帰属における  $c$  の原因としての妥当性や、C. S. Peirce の言うアブダクション (abduction) における説明としての確らしさを表していると捉えることも可能である。現実的場面において、原因の候補となり得る多数の要因の中から、できるだけ効率的に妥当性の高い原因候補を絞る必要があるとき、 $P(c|e)$  の値を用いることは、やはり適応的ヒューリスティックスと言えらるのだろう。

このようなヒューリスティックスに関連する問題として、Kareev (2000) は、人間は相関の検出の際、 $7 \pm 2$  個程度の小標本から判断しており、そのことがむしろ効率のよい相関の検出に役立っていると論じた。刺激 CI1 の結果が曖昧になった理由として、§4.2 ではノイズの多さを指摘したが、この刺激は総事象 (試行) 数の点で特殊であった点にも注意する必要がある。他の刺激は 10-40 の範囲であったのに対し、CI1 だけは 100 試行を要した。もし被験者が、最初の数試行で大方の判断を決め、残りの情報はあまり重視しないとすると、このような刺激の場合、系列の順序が判断を大きく左右することになりかねない。厳密にはこの問題はどの刺激においても内在するが、総試行数が多い刺激においてはその中の少数の刺激、CI1 の場合、特に 1 試行しかない  $\bar{c}\bar{e}$  の系列位置がより大きな問題となる。本実験では、2つのモデルの予測が分かれる刺激を用意する必要性からやむを得なかったとは言え、Kareev (2000) の主張を前提とすれば、この刺激の妥当性に対する疑問の余地は残る。

しかし、総事象数の多さを問題としないモデルもある。それは、連合論的モデル (Rescorla & Wagner, 1972)<sup>5)</sup> や Bayes 的モデル (Anderson, 1990) のように、頻度情報を学習の強度として累積・更新していくタイプのモデルである。そのようなモデルの場合、学習曲線が観察されることを期待しているため、むしろ、本研究ではスコープに入らなかった相対頻度と絶対頻度の違いが問題となる。確率値のみならず、絶対頻度も選択に影響を及ぼすことは、これまでにしばしば指摘されてきた (e.g., Estes, 1976; Denes-Raj & Epstein, 1994; Slovic, Monahan, & MacGregor, 2000)。本実験では、相対頻度が一致する刺激は意識的に避けたため明確な

5) 連合学習の Rescorla-Wagner モデル (Rescorla & Wagner, 1972) は、ある制約のもとで、随伴性モデルに漸近的に一致する (Chapman & Robbins, 1990)。

結論は述べられないが、このようなことが、結果に影響した可能性は否定できない。例えば、CC1とCC2の各事象確率は互いに類似していたが、絶対頻度はCC2が2倍であった。§4.2ではCC2のノイズの多さを指摘したが、絶対頻度の違いがノイズ(いわゆる例外としての $c\bar{e}$ や $\bar{c}e$ の存在)を余計に際立たせた可能性もあるかもしれない。ただ、連合論的モデルが予測する因果判断の学習曲線様の変化については、否定的な実験結果もあり、議論を呼んでいる(e.g., Anderson & Sheu, 1995; 嶋崎, 1999; Lober & Shanks, 2000)。

§4.2, 4.3, 4.4のいずれの分析においても、PCモデルの予測力を完全に否定することはできなかった。このことは、すべての被験者が全く同じ方略を用いているのではなく、実際には両モデルが予測する方略が混在している可能性を示唆している。それと同時に、§4.4の確信度の結果から示唆されたように、一人ですら複数の方略の間で揺れ動いている可能性もある。今後、このような集団や個人における方略の混在状況のモデル化が期待される。また、因果関係は悉無律の関係というよりは、むしろ強さを持った関係概念と考えられるので、その強さを予測できるようモデルを定量化する必要もあるだろう。

本研究では、直感的な因果知覚のモデル、すなわち、強い事前信念を前提としない状況において、人はどのような暫定的な因果帰納をするか、という問題を扱った。しかし、現実世界、特に科学や医学において人間が行っている因果判断は、多くの知識や信念を前提とし、より分析的な過程を含んでいる。例えば、因果と相関の違いを理解するにしても、多くの知識と複雑な分析的過程が必要となる。人間の因果推論は、他の推論過程(Evans, 1984; Evans & Over, 1996; Sloman, 1996; Stanovich, 1999; Denes-Raj & Epstein, 1994)と同様、本研究で扱ったようなヒューリスティック過程の他に、より高度な分析的過程も存在し、2段階から構成されると考えるのが妥当であろう。

実験の結果、「条件」を最頻反応とする刺激は一つもなかったが、このことに積極的含意を見出すことも可能かも知れない。すなわち、一方がもう一方の「条件」であるという関係(条件関係と呼ぶ)の概念の形成のためには、この実験のような2事象間の共変動の観察だけでは不十分かもしれないとい

うことである。もしそうだとすると、条件関係の帰納の説明として次の2つの可能性が考えられる。

1つは、Cheng & Novick (1992)の焦点集合の概念に説明を託すものである。しかし、焦点集合とは、文脈に従って半自動的に定まると同時に、多くの人に共有されるものと考えられるが、今のところ、どのような場合にどのようにして焦点集合が決定するのか形式化されておらず、不明確な点が多い。この点に問題が残されている。ただ、指摘しておくべき重要な点は、現在の文脈と異なる別の焦点集合を想定するというのは、かなり分析的な処理を要すると考えられるので、二段階説における第2段階に相当する処理と考えた方がよさそうだとすることである。そうだとすれば、本研究で扱ったヒューリスティック過程において条件反応が少なかったことも納得がいく。

もう1つは、以下で述べる「条件という概念は原因候補事象が複数存在するときに利用される」とする仮説に基づく説明である。複数の原因候補事象について、Cheng & Novick (1992)は、ANOVAモデルに類似した高階の対比を考えた。しかし、その計算の複雑さは、事象数の増加につれて指数関数的に増大する(計算論的に手に追えない問題となる)ことに留意するべきである。このような方略は、適応的観点から妥当性が極めて低いと言わざるを得ない。因果概念が、外界を認知するためのしくみとして進化の過程で獲得された(戸田, 1986)とすれば、無数の原因候補事象が存在する現実世界において、淘汰圧により適応性を備えたエージェントが、ただか暫定的な結論を出すために、指数関数的にコストが増大する方略を使用するとは考えにくい。

DHモデルの主張は、 $c$ と $e$ が双条件的関係にある場合に因果関係が知覚されるとするものである。ならば、条件関係(=非因果関係)は非双条件的と言えるかもしれない。人は、非双条件的な因果関係を持つ事象をまとめて「条件」とラベル付けすることにより、因果関係の強さの計算が複雑化するのを避けていると考えることはできないだろうか。つまり、条件とは、原因候補事象が複数存在する場合に、計算量を爆発させないための概念装置として機能している可能性がある。以上の仮説が正しいとすると、条件に関するDHモデルの予測は、3つ以上の事象間の共変動情報からの帰納においてこそ検証されるべき価値があると言えるだろう。



## 文 献

- Agresti, A. (1990). *Categorical Data Analysis*. New York: John Wiley & Sons.
- Anderson, J. R. (1990). *The Adaptive Character of Thought*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Anderson, J. R. & Sheu, C.-F. (1995). Causal inferences as perceptual judgments. *Memory & Cognition*, **23** (4), 510-524.
- Bishop, Y. M. M., Fienberg, S. E., & Holland, P. W. (1975). *Discrete Multivariate Analysis: Theory and Practice*. Cambridge, MA: The MIT Press.
- Chapman, G. B. & Robbins, S. J. (1990). Cue Interaction in Human Contingency Judgment. *Memory & Cognition*, **18** (5), 537-545.
- Cheng, P. W. (1997). From Covariation to Causation: A Causal Power Theory. *Psychological Review*, **104** (2), 367-405.
- Cheng, P. W. & Novick, L. R. (1990). A Probabilistic Contrast Model of Causal Induction. *Journal of Personality and Social Psychology*, **58** (4), 545-567.
- Cheng, P. W. & Novick, L. R. (1991). Causes Versus Enabling Conditions. *Cognition*, **40** (1-2), 83-120.
- Cheng, P. W. & Novick, L. R. (1992). Covariation in Natural Causal Induction. *Psychological Review*, **99** (2), 365-382.
- Denes-Raj, V. & Epstein, S. (1994). Conflict Between Intuitive and Rational Processing: When People Behave Against Their Better Judgment. *Journal of Personality and Social Psychology*, **66** (5), 819-829.
- Estes, W. K. (1976). The Cognitive Side of Probability Learning. *Psychological Review*, **83** (1), 37-64.
- Evans, J. St. B. T. (1984). Heuristic and Analytic Processes in Reasoning. *British Journal of Psychology*, **75**, 451-468.
- Evans, J. St. B. T. & Over, D. E. (1996). *Rationality and Reasoning*. Essays in Cognitive Psychology. Hove, UK: Psychology Press.
- Gigerenzer, G., Todd, P. M., & The ABC Research Group, (1999). *Simple Heuristics That Make Us Smart*. New York: Oxford University Press.
- Hart, H. L. A. & Honoré, T. (1959/1985). *Causation in the Law* (Second edition). Oxford, UK: Oxford University Press. (First published 1959).
- Hattori, M. (in press). A Quantitative Model of Optimal Data Selection in Wason's Selection Task. *The Quarterly Journal of Experimental Psychology (A)*.
- Jenkins, H. M. & Ward, W. C. (1965). Judgment of Contingency between Responses and Outcomes. *Psychological Monographs: General and Applied*, **79** (1), 1-17.
- Kareev, Y. (2000). Seven (Indeed, Plus or Minus Two) and the Detection of Correlations. *Psychological Review*, **107** (2), 397-402.
- Kelley, H. H. (1967). Attribution Theory in Social Psychology. In D. Levine (Ed.), *Nebraska Symposium on Motivation 1967*, Vol. 15 of Current Theory and Research in Motivations, 192-238. Lincoln: University of Nebraska Press.
- Lober, K. & Shanks, D. R. (2000). Is Causal Induction Based on Causal Power? Critique of Cheng (1997). *Psychological Review*, **107** (1), 195-212.
- Oaksford, M. & Chater, N. (1994). A Rational Analysis of the Selection Task as Optimal Data Selection. *Psychological Review*, **101** (4), 608-631.
- Rescorla, R. A. (1968). Probability of Shock in the Presence and Absence of CS in Fear Conditioning. *Journal of Comparative and Physiological Psychology*, **66** (1), 1-5.
- Rescorla, R. A. & Wagner, A. R. (1972). A Theory of Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Non-reinforcement. In A.H. Black & W.F. Prokasy (Eds.), *Classical Conditioning II: Current Research and Theory*, chap.3, 64-99. New York: Appleton-Century-Crofts.
- 嶋崎 恒雄 (1999). 随伴性判断の獲得過程に対する連合学習モデルの適用の妥当性に関して. 『心理学研究』, **70** (5), 409-416.
- Sloman, S. A. (1996). The Empirical Case for Two Systems of Reasoning. *Psychological Bulletin*, **119** (1), 3-22.
- Slovic, P., Monahan, J., & MacGregor, D. G. (2000). Violence Risk Assessment and Risk Communication: The Effects of Using Actual Cases, Providing Instruction, and Employing Probability Versus Frequency Formats. *Law and Human Behavior*, **24** (3), 271-296.

- Smedslund, J. (1963). The Concept of Correlation in Adults. *The Scandinavian Journal of Psychology*, 4, 165-173.
- Stanovich, K. E. (1999). *Who Is Rational? Studies of Individual Differences in Reasoning*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- 戸田 正直 (1986). 因果関係の認知について. 『行動計量学』, 14 (1), 60-70.

(Received 24 Mar. 2001)

(Accepted 4 Sep. 2001)



**服部 雅史 (正会員)**

1964 年生まれ. 1990 年北海道大学文学部卒業. 1996 年北海道大学大学院文学研究科博士後期課程単位取得退学. 1997 年立命館大学文学部助教授 (現在に至る). 推論, 問題解決, 意思決定などの高次認知機能に興味を持つ. 特に人間の思考の適応的合理性を明らかにしたいと考えている. 日本心理学会, 日本基礎心理学会, 日本教育工学会, 各会員.