E.T.ジェインズの「確率理論：科学の論理」の紹介（5）

原　宣一

ジェインズの「確率理論：科学の論理」は合理的で常識との定性的一致および首尾一貫していること（デシダレータ）を下にした論理の展開である。それは単純に数学の問題であって、結論はラプラスの確率に不動の基礎を与えるものであった。前回は同じく第２章から和の規則が必然的に導かれることを説明し、三段論法が積の規則に対応するところまで紹介した。今回は同じく第２章から、デシダレータに基づく論理の展開がラプラスの確率定義に一致することを説明する。

数値

我々はこれまでのところ、ロボットの頭脳がある明確な物理的過程を実行することによって操作できるように、実数に関係付けられていなければならないということが受け入れられて、ロボットが妥当さを操作できる最も一般的な首尾一貫規則を見つけた。一方、次の二つの明確な周囲状況はロボットの頭脳を設計する我々の仕事がまだ終わっていないことを示している。

最初に、規則(2.63)、(2.64)は異なった命題の妥当さが如何に各々他に関係付けられていなければならないかに関して、ある制限が置かれる一方で我々はまだ如何なるユニークな規則も持っていないばかりか、ロボットが妥当さの推論を行うことが出来る無限に多くの可能な規則を持っているということである。単調関数 $p(x)$のあらゆる異なった選択に対応して、異なった内容を持った、異なった規則の組があるように見える。

第二に、妥当さの実際の数値が問題の始めにロボットがその計算を始められるように割り当てられるべきだとは、これまでに与えられた何にも我々に語っていない。如何にロボットは背景情報の妥当さの明確な数値への初期符号化を行えるのであろうか。このために我々はまだ使っていない「インタフェース」デシダレータⅢb、(1.39)のⅢc を呼び起こさなければならない。

次の解析はこれら二つの質問に面白くかつ予期せぬ方法で答える。妥当さ$(A_1+A_2+A_3|B)$のために、少なくとも三つの命題（A_1、A_2、A_3）の一つが真であるということを求めよう。我々はこれを以下のように拡張した和の規則(2.66)の二つの適用で見つけることが出来る。最初の適用は、

$$p(A_1+A_2+A_3|B)= p(A_1+A_2|B)+ p(A_3|B)- p(A_1A_3+A_2A_3|B) \qquad (2.80)$$

を与える。ここで、我々は最初に(A_1+A_2)は単一の命題であると考え、次の論理関係式を用いた。

$$(A_1+A_2)A_3= A_1A_3+A_2A_3 \qquad (2.81)$$

(2.66)を再び用いて、我々は次のようにグループ分け出来る７項を得る。

$$p(A_1+A_2+A_3 \mid B) = p(A_1 \mid B) + p(A_2 \mid B) + p(A_3 \mid B) - p(A_1A_2 \mid B) - p(A_2A_3 \mid B) - p(A_3A_1 \mid B) + p(A_1A_2A_3 \mid B) \tag{2.82}$$

さて、これらの命題は相互背反であると想定しよう。即ち、証拠 B はそれらのどの二つも同時に真ではないということを意味する。

$$p(A_iA_j \mid B) = p(A_i \mid B)\delta_{ij} \tag{2.83}$$

そのとき、(2.82)の後の 4 項は消えてしまう。そして我々は次式を得る。

$$p(A_1+A_2+A_3 \mid B) = p(A_1 \mid B) + p(A_2 \mid B) + p(A_3 \mid B) \tag{2.84}$$

より多くの命題 A_4、A_5、等を加えて、我々が n 個の相互背反の命題$\{A_1 \cdots A_n\}$を持つならば(2.84)は次式のように一般化されるということは帰納法により容易に示せる。

$$p(A_1 + \cdots + A_m \mid B) = \sum_{i=1}^{m} p(A_i \mid B)、\ 1 \le m \le n \tag{2.85}$$

これは我々がこれから定常的に用いる一つの規則である。

慣例的な提示で,(2.85)式は通常基本的なものとして導入されるが、それはその理論の任意な一つの公理としての扱いである。現在のアプローチはこの規則が単純な首尾一貫性の定性的条件から演繹可能なものであるということを示している。(2.85)を原始的で基本的な関係として見る視点は我々が特に避けたいことである。

さて、命題$\{A_1 \cdots A_n\}$は相互背反であるばかりか、網羅的であると想定しよう。即ち、背景情報Bはそれらのただ一つが真でなければならないということを鼓舞する。その場合、和(2.85)は m=n に対して 1 にならなければならない。

$$\sum_{i=1}^{n} p(A_i \mid B) = 1 \tag{2.86}$$

これ一つでは個々の数値 $p(A_i \mid B)$ を決めるのに十分ではない。情報Bのさらなる詳細に依存して、多くの異なった選択が適切であったに違いない。そして、一般にBの論理的解析により $p(A_i \mid B)$ を見つけることは困難な問題である。事実それは、Bに含まれているかも知れない複雑な情報の多様さに終わりがないので終わりの開いた問題である。そして、それ故にその情報を $p(A_i \mid B)$ の数値に通訳する複雑な数学の問題に終わりがない。これが最も重要な現在の研究問題の一つである。情報Bを $p(A_i \mid B)$ の数値に通訳するために我々が発見できるあら

ゆる新しい原理がこの理論の応用に役立つ。

二つの異なった問題を考えてみよう。問題Ⅰはちょうど形式化されたものの一つである、我々は一組の相互背反で網羅的な命題$\{A_1 \cdots A_n\}$を与えられ、$p(A_i B)_I$を評価することを求める。問題Ⅱは二つの命題 A_1、A_2 のラベルが交換されているという違いがある。これらのラベルはもちろん全く任意である。どちらの命題を A_1 と呼び、どちらを A_2 と呼ぶかは差がない。問題Ⅱにおいて、それ故、次のように与えられた、我々は一組の相互背反で網羅的な命題$\{A'_1 \cdots A'_n\}$を持つ。

$$A'_1 = A_2 \qquad (2.87)$$
$$A'_2 = A_1$$
$$A'_k = A_k、\qquad 3 \le k \le n$$

そして、我々は量$p(A_i B)_{II}$、i=1,2,・・・、n、を評価することを求める。

ラベルを交換することにおいて、我々は異なってはいるが近い問題を生成した。問題Ⅰにおいてロボットが持っていたA_1の知識についての状態がどんなものであれ、問題ⅡにおいてA'_2について、それらは同じ命題なのであるから、同じ知識の状態でなければならないということは明白である。与えられた情報Bは二つの問題で同じである。両問題で命題$\{A'_1 \cdots A'_n\}$の同じ合計を意図している。それ故、

$$p(A_1 | B) = p(A'_2 | B)_{II} \qquad (2.88)$$

同様に、

$$p(A_2 | B) = p(A'_1 | B)_{II} \qquad (2.89)$$

我々はこれらを変形方程式と呼ぶであろう。それらは如何に二つの問題が各々他に関係しているかということのみ記述している。そして、特に命題A_1、A_2が如何に問題Ⅰでロボットに妥当であろうとなかろうと、どのように見えようとも、それ故に情報Bがどんなものであろうと成立しなければならない。

しかし、今や情報Bが命題A_1とA_2の間で無関係であると想定しよう。即ち、もしそれが一つについて何かを言うならば、他についても同じ事を言っており、ロボットに一方を好むいかなる理由をも与えるものを含んでいない。この場合、問題ⅠとⅡは単に関係していないばかりか、全く等しいのである。即ち、ロボットは問題Ⅱにおける命題$\{A'_1 \cdots A'_n\}$の組について、それらのラベルの付け方も

含めて、問題Ⅰにおける$\{A_1 \cdots A_n\}$ の組についてと同じように、正確に同じ知識の状態なのである。

さて、我々は(1.39)でⅢcにおける首尾一貫性のデシダレータムを呼び起こす。これは知識の同じ状態が等しい妥当さの割り当てによって表されなければならないことを述べている。式でこの叙述は、我々が対称方程式と呼ぶものである。

$$p(A_i \mid B)_I = p(A'_i \mid B)_{II}, i = 1,2,\cdots,n \tag{2.90}$$

しかし今、(2.88)、(2.89)、(2.90)を結合して、我々は次式を得る。

$$p(A_1 \mid B)_I = p(A_2 \mid B)_{II} \tag{2.91}$$

換言すると、命題 A_1 と A_2 は問題Ⅰにおいて等しい妥当さを割り当てられなければならないのである。そして、もちろん問題Ⅱにおいても。

この点で、この主題においてあなたの個性や背景に依存して、結果（2.91）により、あなたは非常に感銘を受けるか、非常に失望するかのどちらかであろう。我々がちょうど与えられた議論は妥当さを割り当てるためのグループ対象性原理の「赤ん坊」版なのである。それは「事前非通知」を割り当てる一般問題として後の章において大きく拡張される。

もっと一般的に、$\{A''_1 \cdots A''_n\}$を$\{A_1 \cdots A_n\}$ のいかなる順列であるとし、$p(A''_i \mid B)$を決めるということを問題Ⅲとしよう。もしその順列が $A''_k \equiv A_i$ のようであれば、次形式のｎの変形方程式があるであろう。

$$p(A_i \mid B)_I \equiv p(A''_k \mid B)_{III} \tag{2.92}$$

これは如何に問題ⅠとⅢが各々他に関係しているかを示している。そしてこれらの関係は情報Ｂがどんなものであっても成立するであろう。

しかし、もし情報Ｂが全ての命題 A_i の間で無関係ならば、そのときロボットは正確に問題Ⅲにおける命題$\{A''_1 \cdots A''_n\}$の組について、それが問題Ⅰにおいて$\{A_1 \cdots A_n\}$の組がそうであるように、同じ知識の状態にある。そして、再び我々の首尾一貫性のデシダレータはそれが等しい知識の状態に等しい妥当さを割り当てるようｎの対象条件を導いて要求する。

$$p(A_k \mid B) = p(A''_k \mid B)_{III}, \quad k = 1,2,\cdots,n \tag{2.93}$$

(2.92)、(2.93)から、次の式を得る。

$$p(A_i \mid B)_I \equiv p(A_k \mid B)_I \tag{2.94}$$

さて、これらの関係式は問題Ⅲを定義するのに我々が用いた特別な順列がどのようなものであっても成立しなければならない。ｎ!の順列があるから、実際に

n!の等価な問題がある。ここで、与えられたiに対し、指標kは(2.94)における(n-1)の全ての他のものに及ぶであろう。それ故、ただ一つの可能性はすべての$p(A_i \mid B)_I$が等しいということである。実際、既にもしそれが位nの巡回的なものならば単一の順列の考慮によって要求されている。$\{A_1 \cdots A_n\}$は網羅的であるから、(2.86)式は成立し、ただ一つの可能性はそれ故、

$$p(A_i \mid B) = \frac{1}{n}、\qquad (1 \le i \le n) \tag{2.95}$$

で、我々はついに明確な数値に到達したのである。ケインズ(1921)に従い、我々はこの結果を無差別の原理と呼ぶ。

恐らく、読者の直感は回りくどい推論の必要性無しに既にこの結論を導いていたであろう。もしそうなら、少なくともその直感は我々のデシダレータに首尾一貫している。しかし、単に(2.95)を直感的に書き下すことはこの結果の重要性とユニーク性の有難さを人に与えない。

積と和の規則を見つけた後で、単調関数 p(x)のあらゆる異なった選択が異なった規則の組（すなわち、異なった内容の組）に導くだろうから、まだ我々は推論のユニークな規則を何ら見つけていないということが明らかになる。しかし、今や我々はどんな関数p(x)を我々が選ぼうとも、我々は同じ結果(2.95)、そして同じ数値pに導かれるであろう。さらに、ロボットの推論過程は量pの操作によって積と和の規則が示すように、完全に遂行できるであろう。そしてロボットの最終結論はxの項でなくpの項で等しく上手く述べられ得るであろう。

だから、我々は今や関数 p(x)の異なった選択は我々がロボットの内部記憶回路を設計することができたその方法にのみ対応するということが判る。ロボットが推論することになる各々の命題A_iのために、それはA_iの妥当さの程度を表すある数値を貯蔵する一つの記録アドレスを、それが与えられている全てのデータの基礎上で、必要とするであろう。もちろん、数 p_iを貯蔵する代わりにそれはp_iのどの厳密な単調関数をも等しく上手く貯蔵できたであろう。

我々がこれを認識するや否や、p(x)がxの任意な単調関数であると言う代わりに、その点にこれを向けることはもっと多くのことであることは明白である。そして、次のように言える。

妥当さ x≡A|B は(0≦p≦1)で定義されたpの任意な単調関数である。
一つの問題のデータによって硬く固定されたものはpでありxではない。

ユニーク性の質問は、それ故、結果(2.95)によって自動的に処置される。我々のロボットが妥当な推論を行うことが出来る規則の唯一つの首尾一貫した組がある。

我々の妥当な推論の理論は量ｐの項でまったく実行されるということを見て、我々はついにそれらの技術的名称を導入する。今後、我々はこれらの量を確率と呼ぶ。言葉の「確率」はこの時点まで注意深く避けてきた。その言葉は良く知られた口語の意味を持つ一方で、我々にとっては一つの技術用語で、正確な意味を持つべきだからである。しかし、これらの量が問題のデータによってユニークに決められるということが実証されるまで我々は量ｐが何らかの正確な意味を持たされていたと想定する根拠を我々は持っていなかったのである。

今やそれらが妥当さの程度が測られ得る特別な尺度を定義するということが判る。原理的にこの目的に等しく上手く仕える全ての可能な単調関数から、我々はこの特別な一つをより「正確」だからでなく、より便利だから、選ぶ。すなわち、最も簡単な組み合わせの規則、積と和の規則、に従うのは量ｐなのである。この故に、ｐの数値は我々の情報から直接に決められる。

この状況は熱力学における温度と似ている。ここでは全ての可能な経験的温度尺度ｔから、我々は最終的にケルビン尺度 T を用いる決定をする。それがより「正確な」のではなく、より便利だからである。すなわち、熱力学の法則はこの特別な尺度の用語で、それらの最も単純な形式［dU=TdS－PdV、dG=－SdT+VdP、など］を取る。この故に、ケルビン温度の数値は実験で、水や水銀のような特別な物質の特性に独立に、直接に測定可能であるという意味において「硬く固定した」ものなのである。

（５完）