ラプラスによる確率の定義に対する一考察

平成 13 年６月１５日

原　宣一

概要

本稿は命題の真偽に対する確かさの度合いを確率の定義とする立場で、ラプラスによる確率の定義で唯一批判の対象とされてきた「等しく起こりやすい」との条件を検討したものである。この結果、ラプラスの定義は確信の度合いの定義を補足するものであるとともに、確信の度合いの定義に統合して良いことが判った。

まえがき

ラプラスによる確率の定義は物理学では広く採用されているにも拘わらず、工学の分野ではフォン・ミーゼスによる頻度概念の定義に取って代わられている。これまで、一度の事象の生起に対する確かさに関しての表現も、繰り返し観察できる事象を想定した確率によらざるを得ないとされてきた。即ち、確率はその対象に属する物性値のようなもので、統計を取ることによってその値、即ち確率、を推定できるという立場が取られてきた。

ラプラスの定義を古典的な定義と古いものにしてしまったフォン・ミーゼス流の定義は繰り返し観察出来る現象を前提としている。このため、工学分野での実応用に際して、一度の現象に対する確かさを論ずる局面ではしばしば苦しい説明にならざるを得ない状況がある。

例えば、衛星ミッションが成功するかどうかについて思索するとき、その確かさとは思索する人間の心の状態である。衛星ミッション達成確率というものがフォン・ミーゼス流に物性値としてあるものとし、これを推定するものとすると、推定した確率（物性値）の確かさ（心の状態）ついても何らかの表現が必要になる。これまで、ネイマン・ピアソンの仮説検定論と同様な考え方で、前者には信頼度と言う言葉で、後者には信頼水準と言う言葉で区別し、「信頼度 0.9 以上であることを信頼水準６０％で言える」との表現が取られてきた。この表現は極めて巧妙であり、論理に誤りは無い。しかし、当初の衛星ミッションが成功するかについての確かさがどのくらいかという問いかけには直接的な回答にはなっていない。信頼度と信頼水準の数値表現の組み合わせは任意であり、これまで信頼水準の数値は慣習により天下りに与えられてきたに過ぎない。

工学の場合は確かさの問題に対する確率は意思決定への応用である。一度の事象の生起に対する確かさを扱うためには、サベイジによる「命題が真であることの確かさの度合いを確率の定義とする」ことが素直である。

1.　情報が無い場合の確率（二値の場合）

結果についての表現が最も簡単で考察しやすい実験を考える。これは、結果事象として取りうる状態が二通りしか無い実験で、ラプラスによる確率の定義で特別な場合と考えられる。

（強い確信）

この範疇に入る実験例として代表的なものにコイン投げがある。投げられたコインが取りうる状態は表が出るか、裏が出るかの二通りしかない。実世間では一回のコイン投げの表か裏かの結果によって何らかのアクションを強いられることもある。このような時に表がでることに対する確信の度合いが問題となる。コイン投げの実験では、通常、いかさまコインは無いものとの先入観に強く支配され、先験的に表が出る確率はほぼ１／２に近いと見なしてしまうであろう。

（弱い確信）

同じく取りうる状態が二通りしかない実験で、先入観が入りにくいものもある。この代表例として、革袋に入れた碁石を１個取り出す実験を考える。この実験では革袋に入れる碁石は白と黒をどのように入れたか全く判らないものとする。即ち、同数入れたかも知れないし、白石だけかも知れないし、適当に混ぜて入れたのかも知れないとの条件である。

このような実験で、１個取り出す碁石が白であることの確信の度合いを我々の問題とする。

さて、白石が出る確率Ｐ（白）＝ｐとすると黒石が出る確率Ｐ（黒）は１－ｐである。石が出る状態として白か黒かの二通りしかないからである。また、石を入れるところを見ていないので、どちらが出やすいという情報もない。従ってＰ（白）＝Ｐ（黒）でなければならない。Ｐ（白）≠Ｐ（黒）ならばどちらかが出やすいことになるからである。

ｐ＝１－ｐ

故に、ｐ＝１／２となる。

（強い確信）

一方、白色と黒色を同数入れたのを見ていた場合、白石がでることと黒石がでることが等しく起こり易いことである。等しく起こり易いことから、

Ｐ（白）＝Ｐ（黒）

従って、上記と同様にｐ＝１／２となる。

石を入れた状況を見ていなくても、取り出した石が１０００回にも及びその約半数が白石であったという十分な数の実験結果を見た後では、数回連続して白石が出たところでｐ＝１／２に近いとの確信はゆらぐものではない。この場合もほぼ強い確信がもてる状況である。

結果としての事実は、実際に取り出した石がどちらの石であったかということだけである。これに対して確信の度合いとは結果に対する推定であって、実際に結果がでていようといまいと関係なく、その事実を見たか否かで確信が変わる。つまり、確信の度合いは時間軸には関係せず、情報を得ているか否かのみ関係する。

（中間の確信）

同数の石を袋に入れたという伝聞情報のみ得られたのであれば、確信の度合いも少し疑わしくなる。また、手品師が入れたのであれば、自分の目の前で見たことに対しても、自分の目を疑うべきかもしれない。実験結果の個数が不十分な場合も強い確信には至らない。等しく起こりやすいということが、積極的に言える場合でも、それが間違った情報に基づく結論、または勘違いであった可能性も絶対的には否定出来ない。

ところが、上に見たように逆にまったく起こり易さについての情報が無い場合も、確信の度合いとして与えるべき数値は同じ 0.5 なのである。

2.　　情報が無い場合の確率（ラプラスの定義）

ラプラスは確率を次のように定義した。

「実験Ｅの結果としてＮ通りの状態があって、その内で望む状態ＡがＮａとおりあるとき、どの状態も等しく起こり得るとき、Ａが起こる確率Ｐ（Ａ）は次式で与えられる。

$$P(A) = \frac{Na}{N}$$

」

望む状態Ａを考えないときは、ラプラスの定義は次のように書き換えられる。

「実験Ｅの結果としてＮ通りの状態があって、どの状態も等しく起こり得るときどの状態が起こる確率Ｐ（Ai）も次式で与えられる。

$$P(Ai) = \frac{1}{N}$$

」

P(A1)、PA(2)、・・・P(An)　が等しく１／Ｎの確率であると「先験的に」与えられ

る。等しく起こりうるということが判れば、全体が１であることから個々の確率は当然１／Ｎとなる。

しかし、２値の場合と同様、取りうる状態の数がＮ通りあるということだけで「等しく起こり易い」ということが判らない場合も、確率は全部１／Ｎでなければならないのである。もし、一つでも確率が他と異なるものがあればそれは「判らない」ということに反する。従って、全部同じ確率でなければならない。１／Ｎの確率は「先験的に」と言われているが、どの状態が起こりやすいかかの情報が無いときに「必然的に」決まるものなのである。

さいころは普通に作られたものであれば、目のでる場合が６通りあるのでどの目も等しく起こり易いと考えられる。しかし、いかさまさいころであることが判っていても、どの目に細工がしてあるか判らなければ、全ての目が出る確率は等しく１／６であるとせざるを得ない。

結局、「等しく起こり易い」条件は、「情報が何もない」条件と同じ確率になる。ラプラスの定義は先験確率と呼ばれるが、「情報がない」ことと同義なのである。そのとき個々の取りうる状態に対して、その状態になる確率は等しく状態の総数の逆数としなければならないのである。もし、そうしないと情報がないことに反するからである。

数学は抽象上の概念として「同じ」がある。しかし、実世界における「同じ」ということは「違いが判らない」ということである。如何に同じ工程で作った二つのコインであっても微視的には必ず結晶欠陥の有無が相互に異なっている。まして、コインを投げる力のかけ方などが完全に毎回同じであることはあり得ない。人間の能力において違いが判らなければ同じと見なしているに過ぎない。

さらに、問題とする事項として表がでるか裏が出るかであれば、コインの種類が違っていても同じ実験と見なせるのである。

つまり、「同じ」とは「対象とする命題に関して違いがわからない」ということを言い換えたことなのである。

このことからも強い確信として「等しく起こりやすい」と言える状況は、本来「違いが判らない」状況に加えて、先入観として「等しい」ということが入っているものである。いかさまの無いさいころは作った時点で先入観として「等しい筈」ということが入っている。コインも作った時点で表と裏があるということが先入観として入っている。両面とも表が鋳造されたなどとは考えないだけである。

ラプラスの定義における「等しく起こりやすい」の意味は、本来実世界においては先入観なしに強い確信でそのように言えることない。弱い確信で「等しく起こりやすい」とみなす状況しかない。

確信の度合いを確率の定義とすることは情報のない最初の状態では、事前確

率としてラプラスの定義を採用する必然性がある。つまりラプラスの定義は確率を命題の真偽に関する確信の度合いとする定義に含まれている。

3.　確信密度

積極的に等しく起こり易い事が判っているときと、情報が全くないときとで何がことなるのであろうか。確信の度合いを確率と考えるとその確率密度分布が違うのである。ここで確信の度合いを確率と定義したことにより確率密度の代わりに確信密度と呼ぶ。

コイン投げの場合で表すと、どちらの状態が生起しやすいかについて強い確信として等しく起こり易い時は、「表がでる」との命題の確かさｐに対して、ｐに対する確信密度分布$\pi(p)$が

$$\pi(p)=\delta(p-1/2)$$

で表すことが適切な状態である。ここでδはディラックのデルタ関数である。

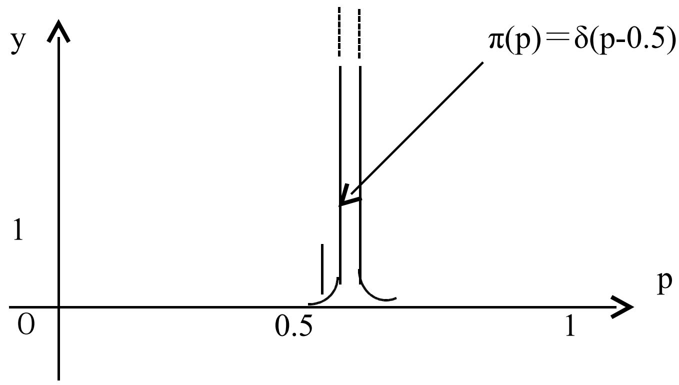

図１　ｐに対する強い確信密度

一方、情報が全く無い場合はｐについて０から１までの一様分布で表すことが適切である。ｐについては特に確率であるということしか情報が無いからである。

$$\pi(p)=1 \qquad 0 \leqq p \leqq 1$$

$$\pi(p)=0 \qquad p \leqq 0,\ 1 \leqq p$$

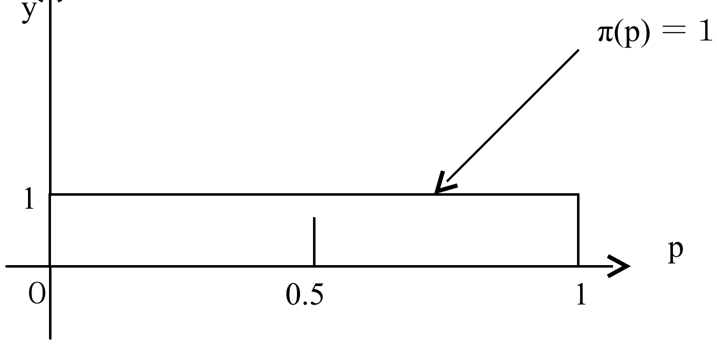

図２　ｐに対する弱い確信密度

情報が無いときも、等しく起こりうる時も確率は１／２であったことから、確率は確信密度の期待値を取ることが適切である。

上述の２例は極端な場合で、通常は何らかの情報がある場合がある。この時の確信密度は上述の２例の中間としてある関数形になると考えられよう。

$$\pi(p) = f(p)$$

この場合の確率は、　　　　ｐ $= E(f(p))$

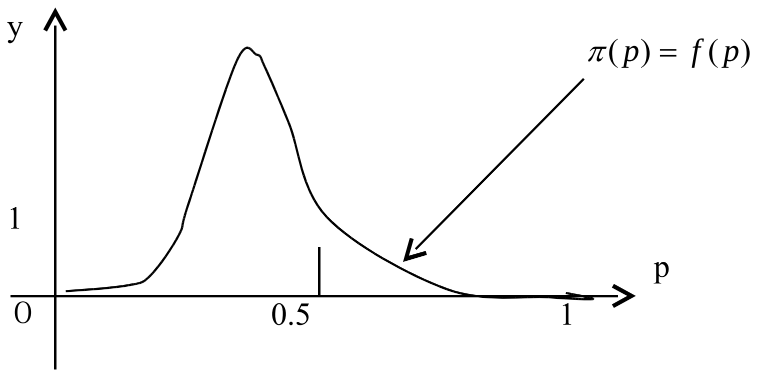

図２　ｐに対する通常の確信密度

何も情報がない状態から、情報を得ることによって確信密度が変わって行く。事前密度が情報を得た後に得られる事後密度はベイズの定理で関係づけられている。

ベイズの定理は、事後密度は事前密度にデータの尤度を乗じたものに比例するということを述べている。

$$\pi(p \mid x) \propto \pi(p) \cdot L(x \mid p)$$

ラプラスの確率の定義では、結局取りうる状態の数を識別することが確率を求めることの主要部分となる。すべての状態の内で望む状態はどれかが決まればそれに当てはまる状態を数えるだけで確率が求まる。

ラプラスの師であったダランベールは２個のコイン同時投げで、取りうる状態は、２個とも表、１個表で１個裏、２個とも裏、の３通りあるので、それぞれの状態に１／３の確率であるとした。ラプラスはこの３つの状態は等しく取りうるのでなく、コインはそれぞれを区別して、１／４、１／２、１／４とすべきであることを示した。実際に多くの実験して相対頻度の変遷をみるとラプラスの与えた確率に近くなることがわかる。しかし、ダランベールが間違っていたわけではない。状態の総数が３通りしかないと考えられる状況であったらそうするしかない。数回の試行の結果を見てベイズの定理で事後の確信密度を求めてその

期待値を求める作業を繰り返すと得られた確率はラプラスの与えた確率に近づく。

この例から、取りうる状態の数は根元事象を数えるべきであることが判る。

結語

確率は命題の真偽に対する確信の度合いと定義することが、実際の応用に自然である。ラプラスの定義で非難のあった「等しく起こりやすい」という条件は先入観の無い場合は情報の無いときと等価であり必然的にラプラスの定義によらざるを得ないことを示した。また確信の度合いの定義を採用する時にラプラスの定義は必然的に組み込まれるべきものであることを示した。

（付録－１）

情報理論では生起する事象がｎ通りある時にどの事象が出るかに関してどの事象についても確信の度合いは 1/n であるとするとき情報量が最小であることが示される。

「仮説 $H_1, H_2, \cdots, H_n$ は互いに排反、かつ完全であり、これらのおのおのに対する信念の度合いを $\pi_1, \pi_2, \cdots, \pi_n$ とすると

$$I = \sum_{1}^{n} \pi_i \ln \pi_i$$

を H_i に関する情報量であると定義する。」

（付録－２）

［ベイズの定理］

ｍのもとでのａの条件付き確率$P(a\,|\,m)$とかくと、

$$P(a\,|\,m) = \frac{P(am)}{P(m)}$$

$$P(a_i b) = P(a_i\,|\,b)P(b)$$
$$= P(b\,|\,a_i)P(a_i)$$

$$\therefore\quad P(a_i\,|\,b) = \frac{P(b\,|\,a_i)P(a_i)}{P(b)}$$

$$P(b) = P(b\,|\,a_1)P(a_1) + P(b\,|\,a_2)P(a_2) + \cdots + P(b\,|a_n)P(a_n)$$

$$P(a_i\,|\,b) = \frac{P(b\,|\,a_i)P(a_i)}{P(b\,|\,a_1)P(a_1) + P(b\,|\,a_2)P(a_2) + \cdots + P(b\,|a_n)P(a_n)}$$

$P(a_i\,|\,b)$ ：事後確率

$P(a_i)$：事前確率

$P(b\,|\,a_i)$ ：ｂの尤度