E.T.ジェインズの「確率理論：科学の論理」の紹介（７）

原　宣一

合理的で、常識との定性的一致し、首尾一貫していることを条件に論理を展開すると、必然的にラプラスの確率を用いることになる。そして、情報理論で基礎においたシャノンの条件が､本質的にジェインズのデシダレータと同じであった。この結果、一般的に確率は最大エントロピー原理で決められるべきことが判った。

ラプラスの確率が正統派統計学者達によって古典確率として追われた最大の理由はその定義にある「同等と見られるならば（equally likely）」という表現に批判があったのだが、ラプラスが導いた連続則が誤解されたことも理由であった。

ラプラスの連続則

ラプラスの連続則とはそれぞれが独立な n 回の試行で、 r 回成功したとの情報のみ与えられたとき、次の試行で成功する確率 P は次式になるというものである。

$$P = \frac{r+1}{n+2} \tag{7.1}$$

ラプラスの連続則で $n=0$ のときは必然的に r も０であり、確率は 0.5 となる。つまり情報が何も無いときは五分五分である。そして n が大きくなれば確率は相対頻度 r/n に近くなる。この式は色々な場面で実感を非常に良く説明するものである。

ラプラス自身はこの式の説明のため「明日また太陽が昇る確率」に応用し、逆に誤解を生んでしまった曰くつきの式であるが、ジェインズは確率を与える基本的な式であると言っている。

さて、ラプラスの連続則（7.1）は筆者にとっても、因縁のある式である。今から 30 年以上昔になるが、データ数が少ないときの確率の決め方が無いものか考えていた。当時の一般的な方法は、確率がフォン・ミーゼスによる相対頻度の極限値という定義であったため、データは少なければ何も言えないということでしかなかった。30 個以上のデータが必要とされていた。筆者は 20 個なら何故駄目か、10 個しかなくても何か言えるのではないか、1 回か 2 回の試験では全く何も言えないのかを考えていた。正統派統計学の教科書では当然のことながら、平均や分散の計算も出来ないほどデータ数が少なくてはどうしようもないということである。特に、データ数が少ないのに分布形状を仮定して良いものか心配であった。

そこで、本屋を回って確率統計の本を手に取り、ベイズ流統計学があることを知った。特に、これだと思って読んだのが竹内・新家共訳の D.V.リンドレー著「確率統計入門１」で副題に小さく＜ベイズの方法による＞と書かれていた。そして、筆者もベイズ流統計学に従って、(7.1)を導いたが、ラプラスの連続則であったことは知らなかった。しかし、あまりにも簡潔な式であったため誰かが誘導しているに違いないと思い放置してあった。最近になって、この式を日本信頼性学会に発表することを思い立ち、発表を申し込んだ。その直後に、ブルーノ・デ・フィネッティの教科書“ Theory of Probability”に、この式がラプラスの連続則として紹介されていることを知った。ラプラスの証明方法をインターネットで調べていてジェインズに行き当たったのである。

ジェインズの「確率理論：科学の論理」には第６章と第 18 章にこの式の誘導が示されている。ラプラスはさらにこの式を一般化している。第 6 章の壷モデルからの誘導も第 18 章のラプラス自身による誘導も長くてかつ難解である。そこで、筆者による方法は論理の展開が厳密さを欠いていると思われるが、結果の式は同じだし導出が簡単なので、ほぼ昔の表現のままでこれを説明する。ジェインズのロボットに比較して、筆者は知能的ロボットを意識したアンドロイドを登場させている。なお、第６章にはあっと驚く解釈も載せられている。

まえがき

工学分野ではこれまで１回だけしか問題にしていない事柄の生起の確率でも、多数回の観察が可能ということを前提にした確率の考え方によって決める以外に客観的に正確な方法はないとされてきた。

観察できる回数の多寡に関わらず事象の生起についての確からしさが知りたいことである。そして元来確率の意味するものは確からしさという主観的な観念を数値化したものである。すなわち心の問題である。実際、多くの観察結果を見れば誰もが確からしさを深めるものであり、観察数が少なければ少ないなりに確かさの程度はある筈である。

人間の心の問題を客観的に扱うことが可能かについては、議論のあるところなので、ここではデータを見ることによって確信の度合い決めるアンドロイド（人造人間）を想定する。アンドロイドが人間と違うところは、合理的に決められた規則に従って確信の度合いを答えることである。

最も簡単な事例として、結果が合格か不合格か（go/no-go）のみ得られる属性試験の場合にアンドロイドがどのようにして確信の度合いを持つか、即ち、合理的に決めた規則はどうあるべきかについて考察する。

属性試験

ある種の試験は結果が合格か不合格かのどちらかしか判らない。多くの火工品は作動するか否かのこの種の試験しか出来ない。大量生産が出来て、何個も試験できる場合は、そうすることにより十分な確信の度合いを得ることが出来るが、高価格の故に少量生産であれば、試験に供しうる個数も限られてくる。ある品物をどこからか手に入れて１個も試験せずにその品目の作動を確信できるであろうか。試験個数が１個であった場合はそうであろうか。また、数個あった場合はどうであろうか。

通常は購入先の信用などから判断するがこのような情報が全くない場合を考える。アンドロイドは生まれたばかりで購入先の信用度などは教えられていない。ただ、試験個数とその結果の情報だけを得た後で、アンドロイドはどのような確信の度合いを持つことになるのかを求めることが本論の課題である。

確信の度合いとしての確率

工学分野で通常採用されているフォン・ミーゼスによる頻

度概念の確率の定義とは異なるサベイジの確信の度合い（Degree Of Belief）としての確率の定義を採用する。

確率 p とは「命題に対する確信の度合いである。」即ち、命題が真であることの確からしさについてその度合いを数値で表現したものである。

このように定義した確率も、コルモゴロフの確率の公理を満たすことが出来るので、すべての数学的確率論の定理が使える。

$p=1$ ：命題が絶対的に真であるとの極めて強い確信を表す。
$p=0$ ：命題が絶対的に偽であるとの極めて強い確信を表す。
$p=0.5$ ：命題の真偽が全く判らない時の確信の程度を表す。

今、白石と黒石を入れた革袋から１個石を取り出す実験を考える。命題は「取り出した石は白である。」この命題が真であることに対して抱く確信の度合い、即ち確率 p はどのようなものかを問題とする。

革袋に石を入れところを見ていて、それが白石だけであったということを知っていたら、 $p=1$ である。黒石だけであったことを見ていたら $p=0$ である。黒３０個と白７０個を入れたことを知っていたら、$p=0.7$ である。白と黒を同数入れたことを知っていれば $p=0.5$ である。これらはラプラスの定義による確率であると言っても良い。

これらは p の値については強い確信がある場合であるが、石を入れるところを見ていない場合は p の値についても確信がない状況となる。全部白石かもしれないし、全部黒石かも知れない。判っていることは、白石か黒石のどちらかであるということだけだからである。

このような状況の場合、本命題が真である確かさの度合い、即ち、確率 p は、白の場合と黒の場合の二通りしかないので、やはり $p=0.5$ となる。白、黒同数の石を入れたのを見ていた場合も、同じ $p=0.5$ である。両者の違いは p に対する確信の度合いの内容、つまり密度分布 $\pi(p)$ の形状が異なるのである。

上述のことを適合させるためには、「確率に対する確信密度の期待値はその確率に等しい。」とすることが必要である。これにより、確信密度の分布に関わらず、ラプラスの確率定義を常に満足させることが出来る。（この部分の論理が少しあいまいであった。）

$$p=E\{\pi(p)\} \tag{7.2}$$

前者は石を入れたのを見ていたという状況で、p の値について強い確信を持つことが出来る場合であり、その確信密度分布はディラックのデルタ関数で表すことが出来る。

白石と黒石を同数入れたことを見ている場合は、

$$\pi(p)=\delta(p-0.5) \tag{7.3}$$

逆に、後者の場合は石をどのように入れたのか全く見ていないので、最も弱い確信の度合いの状態であり、その密度分布は０から１までの等分布で表すことができる。

$$\pi(p)=1 \tag{7.4}$$

後者の場合でも、石を１個、２個取り出して結果を見ることにより p についての確信がだんだん強くなる。つまり、データを見ることにより p の値についての確信の度合い密度が変わるのである。

確信密度 $\pi(p)$ は一般に次のような性質を持っている。

$$\begin{aligned}&\pi(p)=0、\quad p<0、\ p>1\\&\pi(p)\geq 0、\quad 0\leq p\leq 1\end{aligned} \tag{7.5}$$

さらに、規格化の条件として次式を加えて良い。

$$\int_{-\infty}^{\infty}\pi(p)dp=\int_{0}^{1}\pi(p)dp=1 \tag{7.6}$$

データを見た後の確信の度合い

デルタ関数で表される程、事前の確信が強ければ、有限のデータを見ても確率は不変である。しかし、事前の確信が弱ければデータを見ることによって確信の度合い、即ち確率が変化する。この状況はベイズの定理が説明している。

ベイズの定理はつぎのように主張している。

「データを見た後の事後密度はデータの尤度と事前密度の積に比例する。」

p に対する事後密度の期待値が求める確率 p である。

属性試験は結果が成功か不成功かだけのデータを提供する試験である。これは成功を白石とみることにより、革袋に入れる石を見ていない場合の実験に相当する。皮袋の中は全部白石かも知れないが黒石がかなり混ざっているかもしれない。（実際の実験は必ずしも独立とは言えないかも知れない。）

まず、初めの状況は確信の度合いである確率 p の事前密度は（7.4）式とすることが前述の考察により妥当である。

データ X は $(x_1,x_2\cdots)$ は成功、不成功の列であるが、順序は関係しないので、試験個数 n と成功個数 r だけが事後密度に関与する。

データ X を見た後の事後密度 $\pi(p\,|\,X)$ は

$$\pi(p\,|\,X)\propto L(X\,|\,p)\,\pi(p) \tag{7.7}$$

ここで $L(p\,|\,X)$ は p に対するデータ X の尤度である。

試験個数 n のうち成功個数 r を得たというデータの尤度は

$p^r(1-p)^{n-r}$ であるから

$$\pi(p\,|\,X)\propto p^r(1-p)^{n-r}\times 1 \tag{7.8}$$

（7.6）式の条件から定数を決めると、

$$\pi(p\,|\,X)=\frac{\Gamma(n+2)}{\Gamma(r+1)\Gamma(n-r+1)}p^r(1-p)^{n-r} \tag{7.9}$$

（7.9）式は β -分布に他ならない。従って、データを見た後の確信の度合い、つまり確率ｐは（１）式により期待値を求めると、

$$p=E\{\pi(p\,|\,X)\}=\frac{r+1}{n+2} \tag{7.10}$$

ここで、n は試験個数で、r は成功数である。(7.10）式が、データを見た後のアンドロイドが示す確信の度合いとしての確率値である。

この式はデータが無い状態を成功と不成功の２個のデータとし、試験データｎとで、データ数を重みとする重み付き平均であると解釈できるのである。（７完）