

ベイジアンの源流

——トーマス・ベイズをめぐって

松原 望

1. トーマス・ベイズその人

時代は18世紀の中葉、清教徒革命、王制復古、名誉革命の動乱の時代は遠く去り、イギリスは、2大政党による議員内閣制の発展と円熟を楽しみ、新大陸植民地を経営して、国運の隆盛を謳歌していた。そのイギリスの南、ケント州の小ぎれいな町タンブリッジ・ウェルズに“シオンの山”(Mount Sion)という公民館があった。そこに長老派(プロテスタント)の専任牧師がおり、その名を、トーマス・ベイズ(Thomas Bayes, 1702-1761)といった。彼はその生活を楽しんでいるようであったが、しかし、彼は聖職者であるからであろうか、独身であった。彼の父、ヨシァア・ベイズも、非国教徒(non-conformist, 英国教会にしたがわない者)としてははじめて、公式に聖職に叙任された者の1人であった。

このベイズこそ、今日「ベイズ統計学」(Bayesian statistics)といわれている統計学の考え方の濫觴をなすものである。

しかし、彼のくわしい生い立ちについてはほとんど何も知られていないといってよい。同時代の一流といわれた、確率論方面の大数学者ド・モアブル(de Moivre, 1667-1754)、ダニエル・ベルヌーイ(D. Bernoulli, 1700-1782)とどのような学問上の交信があったかも明らかでない。数学の

教育をどこでどのようにしてうけたかも不明であるが、ロンドンで文学、言語学、自然科学を学んだことは、一応たしかとされている。1742年、王立協会(Royal Society)のフェローに選ばれているが、その基礎となった業績は、哲学などの形而上学の論文であったようである。

彼の手になる数学上の論文は2点を数えるのみであり、それも彼の死後(1764年)、友人R.プライス(Richard Price, 生命保険の創始者の1人とされている)によって出版されたものである。最初のもは、スターリングの公式 $\ln(x!)$ の発散を論じており、あとのものが他ならぬ有名な「ベイズの定理」(Bayes' theorem, Bayes' rule)といわれているものを含んでいる、そう長くはないエッセイである。そのエッセイの中で、今日のベイズの定理に対応している部分は、「第9命題」だけといってよく、それも、現在のベイズの定理にしたのは、S.ラプラス(S. Laplace, 1749-1827)である。実際、エッセイは難読で真意不明箇所多く、「ベイズの定理」の内容は、実はラプラスに帰すべきであるという諸説も多い。

エッセイは「確率の考え方における、ある問題の解法に関する考察」(“An Essay toward Solving a Problem in the Doctrine of Chances”)と題されている。

2. 「原因」の確率—ベイズ統計学の本質

今日、ベイズ統計学はさまざまな形で主張さ

は、

$$\int_0^1 {}_n C_x p^x (1-p)^{n-x} dp$$

$$= {}_n C_x \cdot \beta(x+1, n-x+1) \quad (\text{ベータ関数})$$

$$= {}_n C_x \cdot \Gamma(x+1) \Gamma(n-x+1) / \Gamma(n-x)$$

$$= 1/(n+1) \quad (x=0, 1, \dots, n) \quad (2.4)$$

なることを洞察していると思われ、この時代としては鋭い直感力といわざるを得ない。事象の起こり方の法則について知識を欠く不知ignoranceの状況では、一様分布を要請するという考え方は、そのまま、S.ラプラスに引きつがれ、今日のベイズ統計学において「理由不十分の原理」(Principle of insufficient reason)といわれる重要な原則をなしている。

ベイズの第9命題を引きついで発展させたのは、S.ラプラスである。ラプラスは「原因の確率」(probabilité des causes)という語をはっきり用いて、ベイズの得た結果に明確な意味解釈を与えた。彼は、あまりにも有名な「確率の解析」(Théorie analytique des probabilités)の第1版(1812)で、 $j=1, 2, \dots, n$ を n 個の原因とし、 p_1, \dots, p_n を当該事象の各原因からの起こり方とするとき、その事象が起こったという条件のもとで、それが原因 j から起因したという確率は、

$$p_j / \sum_{i=1}^n p_i \quad (2.5)$$

で与えられる、と述べている。これは、原因の確率分布が一様であるときであり、第2版(1814)では、 w_1, \dots, w_n があらかじめ事前に有している原因の予想を表わす確率分布—いわゆる、「事前分布」prior distribution—であるとすれば、原因 j の確率は、

$$w_j p_j / \sum_{i=1}^n w_i p_i \quad (\text{「事後分布」}) \quad (2.5)$$

で表わされると、数行つけ加えた。これによって、今日「ベイズの定理」といわれている定理の形式と意味解釈が、最終的に整ったのである。

ついでながら、ラプラスはベイズを引用しなかったので、ベイズの定理の功績は、ラプラスに帰すべきか、(定理としては)要を得なかったベイズ

に帰すべきか、後世の議論は別れることとなっている。

3. ベイズ、ラプラスから近代確率論へ

ベイズの定理は、ラプラス以後、主としてヨーロッパ大陸を中心に一定の程度に受け入れられてきた。しかし、批判も数々加えられたとってよい。R.フィッシャーの批判は、最も有名なものである。

批判の根拠は、確率の「客観説」(objective probability)、特にいわゆる「相対頻度的定義」(relative frequency definition)に求められる。Aという事象の確率を、Aの起こった回数の相対頻度の極限として定義するというものである。つまり $p_A = \lim(n_A/n)$ 。信頼係数 $1-\alpha$ の信頼区間 I は、それがサンプル値という固定値から作られているとしても、“無限回 I を作ったときに $1-\alpha$ の割合で当る計算である”という1つの頻度説的約束(あるいは了解)の上に、概念そのものが成立している。

これに対して、ベイズの定理(2.1)はそのような想定を必要としない。頻度説は、「無限」によって確率の概念を作り出したが、ベイズの定理は、 p そのものがすでに確率的存在となっている。事前分布の概念さえ容認すれば、 p の“ありか”に関する確率言明が定理自体の帰結として、ただちに得られる。これは何といっても大きな魅力である。

考えてみれば、古典確率論がすでに圧倒的に頻度説の中にあつた。ベイズのエッセイの中で、プライスが前書きしているように、確率論の先駆者ド・モアブルの仕事

“If a very great number of trials be made concerning any event, the proportion of the number of times it will fail in those trials, should differ less than by small assigned limits from the proportion of the probability of its failing in one single trial” (372-3)

は、大数の法則、つまり、ある事象の確率と相対頻度の収束関係についての法則を述べており、そこでは確率 p は所与(つまり固定)と仮定されている。(これは、むしろ演繹論理に近いとさえいえる。大陸の合理論—演繹論理—は、デカルトから始まったが、その流れの中にあると考えられる。)つまり、ベイズは、これと逆の問題を解いたのである。

信頼区間の考え方をはじめとして、数理統計学の諸方法における確率の考え方をめぐる論争は、むしろ、確率を必ずしも頻度的に解釈しない、いろいろな考え方を生み出した。「論理確率」(logical probability)を主張した J. M. ケインズ、科学方法論の立場を強調した H. ジェフリーズをはじめとして、F. ラムゼイなどが輩出した。

4. 主観確率と統計的決定理論 ——ベイズ統計学の展開——

主観確率(subjective probability)は、確率の頻度説(客観説)に対して、人間が、対象に与えた「確信の度合」(degree of belief)を基礎におく確率の解釈である。ペルヌーイ試行を例にとれば、頻度説では、尤度関数 $f(x;p) = {}_n C_x p^x (1-p)^{n-x}$ で、サンプルの出方を分析する。しかし、主観説(主観確率)は、 p の(与えられた)事前分布 $w(p)$ 、およびベイズの定理から求められた p の事後分布 $w(p|x)$ が主要な働きをする。たとえば(2.1)、(2.5)は一樣な事前分布を想定した、事後分布である。主観説は、ド・フィネッティ (de Finetti), L. J. サベージ (L. J. Savage) によって、精力的に主張され、H. レイファ (H. Raiffa), R. シュレイファー (R. Schlaifer) によって、数学的にふんばられ、主として経営学方面に応用された。また、D. リンドレー (D. Lindley) は、統計学をベイジアン立場から見た。

しかし、何といっても主観確率が大きな役割をはたすのは、A. ワルド (A. Wald) によって定式化された「統計的決定理論」(Statistical Decision

Theory) と組み合わせられたときであろう。ベイズの定理を用いれば、事物の「原因」に対して与えられた確率が簡単に演算され、このような「不確実性下の意思決定」(Decision-making under uncertainty) の理論に対して、有力な手段を提供するのである。これらの考え方をまとめて、「ベイズ統計学」(Bayesian statistics) と称する。ひとつの実例を述べよう。

例 マーケティング行動と利益分岐点

今期、新規に売り出した商品について、来期新たなマーケティング行動 (a_1 あるいは a_2) を行ないたい。それについては、来期の売り上げ伸び率 $\theta(\%)$ が問題になるが、過去のデータはほとんどない^{*}。テスト・マーケティングをランダムに行ない、一応 $x(\%)$ というデータを得た。 a_1, a_2 にともなう利益を

$$g(\theta, a_1) = \alpha_1 + \beta_1 \theta, \quad g(\theta, a_2) = \alpha_2 + \beta_2 \theta$$

とし(ただし、 $\beta_1 < \beta_2$)、また、情報 x は θ を中心に、ばらつき σ^2 (既知) 程度の精確さの正規分布をすることはわかっているとす。 (x は θ の不偏な推定量、 σ^2 は x の抽出誤差を表わす。) このような不確実性下の状況で、 x にもとづいて a を最適に決定する。すなわち χ = 標本空間、 $A = \{a_1, a_2\}$ として、最適な $d: \chi \rightarrow A$ を決める。(いわゆる Bayes 解)

まず a_1, a_2 の機会損失は次のようになる。

$$L(\theta, a_1) = \begin{cases} 0 & \theta \leq \theta_b, \\ \gamma(\theta - \theta_b) & \theta > \theta_b, \end{cases}$$

$$L(\theta, a_2) = \begin{cases} \gamma(\theta_b - \theta) & \theta \leq \theta_b, \\ 0 & \theta > \theta_b, \end{cases}$$

ただし、 $\gamma = \beta_2 - \beta_1$ 、 $\theta_b = (\alpha_2 - \alpha_1) / (\beta_1 - \beta_2)$ (分岐点 break-even point)。よって、 a_1, a_2 にともなう平均的損失—これをリスク risk という—は、

$$R(\theta, d) = \begin{cases} \gamma(\theta_b - \theta) \cdot P\{x \in d^{-1}(a_2)\} & \theta \leq \theta_b \\ \gamma(\theta - \theta_b) \cdot P\{x \in d^{-1}(a_1)\} & \theta > \theta_b \end{cases}$$

となるが、これには θ が入っているので、最適な d の決定を行なうことができない。

^{*} 過去のデータが十分にあれば、 θ は推定される。

θ に関する事前情報(事前分布)をまだ用いていない。 x が正規分布にしたがうので、 θ の事前分布は、その数学的表現の整合上、正規分布の形で表現することにする。マーケティング・プロモーターの見通しから、 θ はおよそ0%~60%の間に必ず入るであろうこと、また、確率1/2で30±10%に入るであろう、というのであれば、 $\mu=30$ 、 $0.675\tau=10$ から $\tau=14.81$ などおけば結局、具体的に $N(\mu, \tau^2)$ の形^{*}における(主観確率)。一応、それを利用することにすれば、最適な d は、平均リスク

$$r(d) = (\sqrt{2\pi}\tau)^{-1} \int_{-\infty}^{\infty} \exp\{-(\theta-\mu)^2/(2\tau^2)\} \cdot R(\theta, d) d\theta$$

を最小化する d である。

統計的決定理論の定式化によれば、これらには統一的な解法がある。すなわち、それは、「事後分布による、平均(機会)損失の最小化」である。

一般に $w(\theta|x)$ を事後分布として、 x に対して、

$$\text{Min}_a \int_{-\infty}^{\infty} w(\theta|x) L(\theta, a) d\theta$$

を対応させるものが、最適な d -ベイズ解、ベイズ決定方式-を与える。

そこで、事後分布の形はさておき、

$$\begin{aligned} & \int_{-\infty}^{\infty} L(\theta, a_1) w(\theta|x) d\theta - \int_{-\infty}^{\infty} L(\theta, a_2) w(\theta|x) d\theta \\ &= r \cdot \int_{-\infty}^{\infty} (\theta - \theta_0) w(\theta|x) d\theta \\ &= \varepsilon'(\theta) - \theta_0 \quad (' \text{は事後分布を表わす}) \end{aligned}$$

であるから、結局、見通しによる平均伸び率 $\varepsilon'(\theta)$ と、分岐点 θ_0 の大小関係で決することになる：

$$\varepsilon'(\theta) \leq \theta_0 \text{ なら } a_1, \varepsilon'(\theta) > \theta_0 \text{ なら } a_2$$

したがって、問題は純粹に、ベイズの定理による事後分布の計算に還元される。(「ベイズ」解の由来。) やや計算があった後、事後分布 $w(\theta|x)$ は、

正規分布 $N(\mu', \tau'^2)$ 、ただし、

$$\frac{1}{\tau'^2} = \frac{1}{\tau^2} + \frac{1}{\sigma^2}$$

$$\mu' = \left(\frac{\mu}{\tau^2} + \frac{x}{\sigma^2} \right) / \left(\frac{1}{\tau^2} + \frac{1}{\sigma^2} \right)$$

^{*} いわゆる「自然な共役事前分布」(natural conjugate prior)といわれるものの1つである。

となることがわかる。(これはよく知られた結果である。)したがって、テスト・マーケティング・データ x によって、大小関係

$$\left(\frac{\mu}{\tau^2} + \frac{x}{\sigma^2} \right) / \left(\frac{1}{\tau^2} + \frac{1}{\sigma^2} \right) \geq \frac{\alpha_2 - \alpha_1}{\beta_1 - \beta_2} (\equiv \theta_0)$$

の形で、 a_1, a_2 を決めることになる。ちなみに、限界利益を $\beta_1=1, \beta_2=2$ 、また $\alpha_1=30, \alpha_2=10$ とすると、 $\theta_0=20$ 。さらに、 $\sigma=5, \mu=30, \tau=14.81$ とすると、

$$11.68 \left(\frac{x}{25} + 0.137 \right) \geq 20$$

これによると $x=40(\%)$ 前後が、 a_1, a_2 の切替点となり、マーケティング行動 a_2 は、来期4割の売上げ伸び率(!)を必要とすることになる。これは、相当に厳しい要求であるが、伸び率 $\theta=0$ のとき、利益が a_1 は $\alpha_1=30$ であるのに対し、 a_2 はわずかに $\alpha_2=10$ であるので、その構造が、決定者をして逡巡せしめる(リスク回避)のであろう。ちなみに、 $\alpha_2=25$ である場合は、 $\theta_0=15$ となり、 $x=30(\%)$ 余りあれば、 a_2 となる。売上げ3割増ということは、一般的に考えられぬことではない。

このように将来についてある程度自由に議論ができるようになったのは、事前分布、ベイズの定理が、中途の論理の鎖をつなげてくれているからである。結論自体の、実際の妥当性は別個の議論があるであろう。しかし、とにもかくにも、ここまでくることができるということが、ベイズ統計学の大きな強みなのである。

以上、所論をまとめれば次のようになる。

ベイズ統計学(Bayesian statistics)は、将来に対する「見込み」、「見通し」、「確信」、「信念」、「経験」を主観確率(subjective probability)として積極的にデータ解析にとり込み、人間のさまざまな意思決定をサポートする、新しい統計学である。

- i) 過去のデータだけに依拠する従来の統計学より、適用できる範囲が広い。確率に広い役割を与える。

- ii) 意思決定をする者としての人間の主体性を尊重する。
- iii) 分析のみならず、意思決定までをカバーする。
- iv) 社会、人文現象を主として対象とし、今日のニーズに答えうる。
- v) 多大な計算を必要とせず、パソコンで対話型でできる。

5. Apologetik

社会の(学問)に対する要求 demand と、学問の側からの供給 supply とを考えると、demand 側のほうが進んでいるというのが、今日という時代の基本的特徴である。社会の意識のほうが、鋭敏に時代を先取りしており、学問の側はとまどい、対応に苦慮し、あるいは逆に学問の堂宇に引きこもってしまうというのが、一般的状況ではなからうか。統計学、ORは、数理科学(mathematical science)の二大分野であるが、確実なものはいよいよよくなっていく—これこそ、「確率」の働き場である—世界でもっと大きな役割をはたすことができる。ということは、逆にいえば、われわれの歴史は、ひたすら精密科学(exact science)への憧憬に生きてきた歴史ではなかったか。確定論的世界観(deterministic view)を打破しきっていないのではなからうか。

話はややそれるが、われわれは、確率天気予報を受け入れた。そこに、とまどいや混乱がなかったわけではない。しかし、多くの人々は、これのほうが情報が多いと感じている。気象庁の大英断である。

法学の分野でもベイズ統計学の考え方のとり入れが始まっている。これらのことからわかるように、不確実性を正面から分析するベイズ統計学の役割は今後大きい。

ベイズ統計学も問題がないわけではない。その最大のもは、(事前分布たる)主観確率である。しかしながら、次のことは確かである。主観説的

統計学と客観説的統計学の差は、主観を、定式化の中に正式に—formal analysis—とり入れるかどうかにある。客観説といえども、理論の外部的前提として、主観ないしは前提を、ある場合にはきわめて無意識に—ときには無限定に—とり入れているということである。正規分布という仮定、 N はきわめて大(漸近理論)という仮定、線形性の仮定、 R^2 に対する判断等である。さらに今日の「データ解析」(data analysis)では、ロジット、プロビット分析におけるように、きわめて大胆に—ベイジアンが事前分布を“おそるおそる”おくの—の対し—関数形が仮定される。また、EDP(Exploratory Data Analysis 探索的データ分析)のように、解釈そのものも分析に包含する考え方もある。私は、むしろこれらは原則的には好ましい傾向と考えている。

主観説において、「主観」が「主観」たる故に論議ないしは検証不要とされる独断論、唯我論が容認される傾向があった。これは深く自省すべきことであり、かつて竹内啓氏がされたベイジアン批判はその点でまさに正鵠を射ているというべきである。ベイジアンは、これらの批判から多くを学び、新しい時代に新しい役割をはたし、発展していくことが期待されている。

参考文献

次のものを掲げた。(i) 主として歴史的発展の跡を辿る。(ii) ベイズ統計学、あるいはそれに関連した教科書(他にも、何点かの良書があることをおことわりしておく)を掲げる。(iii) 本稿のために参考にした文献を掲げる。

- [1] Bayes, T. (1764) 1963 *Facsimiles of Two papers by Bayes* New York, Hafner (“An Essay Toward Solving a Problem in the Doctrine of Chances” and “A Letter on Asymptotic Series from Bayes to John Canton”) They appeared in *Philosophical Transactions*, Royal Society of London, Vol. 53
- [2] Blackwell, D and Girshick, M. A. (1954)

- Theory of Games and Statistical Decisions*, John Wiley and Sons
- [3] Box, G. E. P. and Tiao, G. C. (1973) *Bayesian Inference in Statistical Analysis*, Addison-Wesley Publishing Co.
- [4] Chernoff, H. and Moses, L. (1959) *Elementary Decision Theory* John Wiley and Sons (宮沢光一訳「決定理論入門」紀伊国屋)
- [5] De Finetti, B. (1937) *La prévision : Ses lois logiques, ses sources subjectives*. Université de Paris, Institut Henri Poincaré, *Annales* 7 : 1-68
- [6] De Groot, M. H. (1970) *Optimal Statistical Decisions*, McGraw-Hill
- [7] Edwards, W., Lindman, H. and Savage, L. J. (1963) Bayesian Statistical Inference for Psychological Research, *Psychological Review* Vol. 70, No. 3, 193-242
- [8] Freund, J. E. and Williams, F. J. (1972) *Elementary Business Statistics : the modern approach*, 2nd. ed. Prentice-Hall (福場庸, 大沢豊訳「経済経営系のための統計学入門」上, 下, 培風館)
- [9] Ferguson, T. S. (1967) *Mathematical Statistics : A Decision Theoretic Approach*, Academic Press
- [10] Fisher, R. A. (1956) *Statistical Methods and Scientific Inference* Oliver and Boyd (渋谷政昭, 竹内啓訳「統計的方法と科学的推論」岩波書店)
- [11] Good, I. J. (1965) *The Estimation of Probabilities, An Essay on Modern Bayesian Methods*, M. I. T. Press
- [12] Jeffreys, H. (1939) 1961 *Theory of Probability*, 3rd ed. Oxford Univ. Press : Clarendon
- [13] Keynes, J. M. (1921) 1961 *A Treatise on Probability* Macmillan
- [14] Laplace, P. S. (1812) 1820 *Théorie analytique des Probabilités* 3rd ed. Courcier
- [15] Lindley, D. V. (1965) *Introduction to probability and Statistics from a Bayesian Viewpoint* Vol. 1, 2. Cambridge Univ. Press (竹内啓, 新家健精訳「確率統計入門」1, 2 培風館)
- [16] Press, S. J. (1972) *Applied Multivariate Analysis* Holt, Rinehart and Winston
- [17] Raiffa, H. and Schlaifer, R. (1961) *Applied Statistical Decision Theory*, Graduate School of Business Administration, Harvard Univ. Division of Research
- [18] Savage, L. J. (1954) *Foundation of Statistics* John Wiley and Sons
- [19] Schlaifer, R. (1959) *Probability and Statistics for Business Decisions : An Introduction to Managerial Economics under Uncertainty*, McGraw-Hill
- [20] Wald, A. (1950) *Statistical Decision Functions*, John Wiley and Sons
- [21] Weiss, L. (1961) *Statistical Decision Theory* McGraw-Hill (邦訳あり)
- [22] Zellner, A. (1971) *An Introduction to Bayesian Inference in Econometrics*, John Wiley and Sons
- [23] 藤本 熙 (1968) 統計数理の基礎と応用, 日刊工業新聞社
- [24] 宮沢光一 (1971) 情報・決定理論序説 岩波書店
- [25] 藤本 熙, 松原 望 (1976) 決定の数理, 筑摩書房
- [26] 松原 望 (1977) 意思決定の基礎, 朝倉書店
- [27] 鈴木雪夫 (1977) 統計解析, 筑摩書房
- [28] 林 周二 (1973) 統計学講義, 丸善(362-3頁)
- [29] 太田勝造 (1978) 裁判における証明論の基礎, 弘文堂
- [30] 武隈良一 (1958) 数学史, 培風館
- [31] 雨宮正彦 (1981) 降雨確率予報への評価 気象25 巻6号, 日本気象協会
- なお, 本稿においては, *Encyclopedia of Social Sciences*, Collier-Macmillan を参考にした。