

逐次実験計画における得られない観察値と検閲のある観察値の比較について

兵庫県立大学・経営学部 濱田 年男 (Toshio Hamada)
School of Business Administration
University of Hyogo

1 緒言

逐次実験計画問題は、未知パラメータを含む実験を行うことにより、得られた観察値を用いて未知パラメータの事前分布を更新して事後分布を求め、それを次の実験に用いるという過程を繰り返して、一定期間内の総期待損失の最小化、あるいは総期待利益の最大化を目的とする。従来はこの問題の大部分において、観察値は確実に得られるという仮定の下で研究がなされてきた。しかし、実際には実験を行った結果、完全な情報が得られなくても、部分的な情報が得られる場合もある。例えば、需要の分布が未知のパラメータを持つ在庫管理問題において、未知パラメータに関する事前分布を仮定しておいて、新しいデータを用いて事前分布を事後分布に更新し、その結果得られた事後分布を新たに事前分布として用いることにより、逐次的に決定を行っていくような場合には、Scarf [6],[7]において研究されている。しかしある期において期首在庫量以上の需要がある場合には、不足部分に関する情報を失うことがある。このような場合に失った情報も考慮に入れて、意思決定を行っていく場合に、検閲された観察値 (censored observation) という概念が Braden and Freimer [1] によって提案され、Lariviere and Porteus [5] によって在庫管理問題に応用されている。

複数の実験を比較検討して選択し、各段階においていずれの実験を利用するかを決定する問題は、バンディット問題と呼ばれ、十分統計量の概念を導入することにより、動的計画法で定式化されて、種々の性質が得られてきた。特に実験結果が一様分布、指数分布、ガンマ分布の場合には、再帰方程式を解くことができ、コンピュータによる数値計算で、臨界値の値が計算され、数表が得られている。

実験を行った結果、十分な観察値が得られない場合には、その不十分な観察値を意思決定にどのように用いるかという場合に、それを捨て去る場合と、何らかの形で利用していく場合の2通りが考えられる。本研究ではこのような問題を、バンディット問題を用いて比較検討する。

2 観察値が必ず得られる場合

次のような逐次実験計画問題を考える. 2つの実験 a_0, a_1 があり, それらの実験を行った結果は, それぞれパラメータ u_0, u_1 を含む確率密度関数 $f(z|u_0), f(z|u_1)$ を持つ分布から, 観察値を得るものとする. ここで u_1 の値は既知であるが, u_0 の値は未知であり, 事前分布 $g(u_0|x, y)$ を仮定できるものとする. a_0 を行った結果, 観察値 z を得たとき, 事後分布は $g(u_0|\phi(x, y; z), \psi(x, y; z))$ で与えられるものとする. 例えば, $f(z|u_0)$ が区間 $(0, u_0)$ 上の一様分布で, $g(u_0|x, y)$ がパラメータ (x, y) のパレート分布の場合には, $\phi(x, y; z) = \max\{x, z\}$, および $\psi(x, y; z) = y + 1$ であり, また $f(z|u_0)$ がパラメータ u_0 の指数分布であり, $g(u_0|x, y)$ がパラメータ (x, y) のガンマ分布の場合には, $\phi(x, y; z) = x + z$, および $\psi(x, y; z) = y + 1$ である. 逐次的に n 回の実験を行えるものとして, 各回に a_0, a_1 のいずれの実験を行っても良いものとする. 目的は n 回の実験で得られる観察値の和を最大にすることである.

状態は残りの回数 n と x, y により, $(n; x, y)$ で与えられるものとする. 以下においては

$$h(z|x, y) = \int f(z|u)g(u|x, y)du,$$

$$E[r(Z)|x, y] = \int r(z)h(z|x, y)dz,$$

および

$$R(x, y) = \int zh(z|x, y)dz$$

を定義し, 次のような仮定の下で議論を進める.

A_1 : $\phi(x, y; z)$ は x について連続, x および z について非減少, y について非増加であり, また $\psi(x, y; z)$ は y について連続, x および z について非増加, y について非減少である.

A_2 : $x' > x$ のとき $z' > z$ に対して $h(z|x', y)/h(z|x, y) \leq h(z'|x', y)/h(z'|x, y)$ および $y' > y$ のとき $z' > z$ に対して $h(z|x, y)/h(z|x, y') \leq h(z'|x, y)/h(z'|x, y')$

A_3 : $(x, y) \in S$ に対して $R(x, y) > 0$

A_4 : $R(x, y)$ は x と y について連続, x について狭義単調増加, y について狭義単調減少

A_5 : $(x, y) \in S$ に対して $R(x, y) = E[R(\phi(x, y; Z), \psi(x, y; Z))|x, y]$

A_6 : $c > 0$ と $y \in S_y$ に対して $R(x, y) < c$ であるような $x \in S_x$ と $R(x, y) > c$ であるような $x \in S_x$ が存在する.

$F_n(x, y)$ を状態 $(n; x, y)$ のときに最適政策を用いたときに得られる最大期待総利得とする。また $i = 0, 1$ に対して $F_n^i(x, y)$ を状態 $(n; x, y)$ のときに、まず i を行い、以後は最適政策を用いたときに得られる最大期待総利得とする。このとき

$$F_n(x, y) = \max \{F_n^0(x, y), F_n^1(x, y)\}$$

ここに

$$F_n^0(x, y) = E[Z + F_{n-1}(\phi(x, y; Z), \psi(x, y; Z)|x, y]$$

および

$$F_n^1(x, y) = 1 + F_{n-1}(x, y)$$

となる。この最後の式の右辺第 1 項の 1 は、 a_1 を行った結果得られる期待値である。

この問題において実験結果は必ず観察され、もし未知パラメータを含む実験が選択された場合には、得られた観察値を用いて、事前分布は更新され事後分布となり、得られた事後分布は次の決定のための事前分布として用いられる。

この最適解は、残りの回数 n と y に依存する関数 $s_n(y)$ が存在して、「 $x \geq s_n(y)$ ならば a_0 が最適である」で与えられる。

$s_n(y)$ の値は数値計算により求めることができ、一様分布の場合は Hamada [2] において、指数分布やガンマ分布の場合には Hamada [4] において、数表が与えられている。

3 観察値が得られない可能性がある場合

実際には実験を行っても実験結果が得られない場合がある。たとえば新薬が開発されて、それを既に開発されて十分に使われてきた薬との比較を行ったりする場合に、新薬を投与中の患者が、何らかの理由で投与を継続できなくなる場合も考えられる。このような場合には、新薬に関する情報がまったく得られない場合も考えられる。

実験を行った結果を観察できる確率を p とする。ここに $0 < p < 1$ である。もし結果を観察できない場合には、その実験から得られた情報を無視するものとする。このような場合には、Hamada [3] において論じられている。そこにおいては仮定 $A_1 \sim A_6$ の下で議論を進め、最適解の構造を明らかにしている。

この場合には、状態は残りの回数 n と x, y, p により、 $(n; x, y, p)$ で与えられる。この状態で実験を行った結果、確率 p で状態 $(n-1; \phi(x, y, z), \psi(x, y, z); p)$ に推移し、確率 $1-p$ で状態 $(n-1; x, y, p)$ に推移する。

$G_n(x, y, p)$ を状態 $(n; x, y, p)$ のときに最適政策を用いたときに得られる最大期待総利得とする。また $i = 0, 1$ に対して $G_n^i(x, y, p)$ を状態 $(n; x, y, p)$ のときに、まず i を行い、

以後は最適政策を用いたときに得られる最大期待総利得とする。このとき

$$G_n(x, y; p) = \max \{G_n^0(x, y; p), G_n^1(x, y; p)\}$$

ここに

$$G_n^0(x, y; p) = pE[Z + G_{n-1}(\phi(x, y; Z), \psi(x, y; Z); p)|x, y] + (1 - p)G_{n-1}(x, y; p)$$

および

$$G_n^1(x, y; p) = p + G_{n-1}(x, y; p)$$

となる。

この場合には、確率 p で観察値が得られ、その結果として状態は $(n - 1; \phi(x, y; z), \psi(x, y; z); p)$ となり、また確率 $1 - p$ で観察値が得られないので、状態は $(n - 1; x, y; p)$ となる。

この最適解は、残りの回数 n と y, p に依存する関数 $s_n(y; p)$ が存在して、「 $x \geq s_n(y; p)$ ならば a_0 が最適である」で与えられる。 $s_n(y)$ の値は数値計算により求めることができ、一様分布の場合には Hamada [3] に与えられている。

4 部分的な情報を利用する場合

実験の途中において、その実験を中断する場合に、それまでに得られた部分的な情報を利用する考え方として、Braden and Freimer [1] の検閲された観察値 (censored observation) を用いる。事前分布 $g(u_0|x, y)$ が与えられているときに、 a_0 を行った結果、確率変数 Z の実現値が b 以下ならば観察値 z を得るが、 b より大きいときには $Z > b$ ということだけを得るものとする。ここに $b > 0$ とする。また、このとき事後分布は $g(u_0|\phi(x, y; Z \wedge b, b), \psi(x, y; Z \wedge b, b))$ で与えられるものとする。したがって状態 $(n; x, y)$ において、 $Z = z < b$ ならば状態 $(n - 1; \phi(x, y; z, b), \psi(x, y; z, b))$ に推移し、 $Z \geq b$ ならば状態 $(n - 1; \phi(x, y; b, b), \psi(x, y; b, b))$ に推移する。

$$R(x, y; b) = \int_0^b zh(z|x, y)dz + b \int_b^\infty h(z|x, y)dz$$

として次の仮定を行う。

A_1 : $\phi(x, y; z, b)$ は x について連続、 x および z について非減少、 y について非増加であり、また $\psi(x, y; z, b)$ は y について連続、 x および z について非増加、 y について非減少である。

A_2 : $x' > x$ のとき $z' > z$ に対して $h(z|x', y)/h(z|x, y) \leq h(z'|x', y)/h(z'|x, y)$ および $y' > y$ のとき $z' > z$ に対して $h(z|x, y)/h(z|x, y') \leq h(z'|x, y)/h(z'|x, y')$

$A'_3 : (x, y) \in S$ に対して $R(x, y; b) > 0$

$A'_4 : R(x, y; b)$ は x と y について連続, x について狭義単調増加, y について協議単調減少

$A'_5 : (x, y) \in S$ に対して $R(x, y; b) = E[R(\phi(x, y; Z, b), \psi(x, y; Z, b)) | x, y]$

$A'_6 : c > 0$ と $y \in S_y$ に対して $R(x, y; b) < c$ であるような $x \in S_x$ と $R(x, y; b) > c$ であるような $x \in S_x$ が存在する.

状態は残りの回数 n と x, y, b により, $(n; x, y; b)$ で与えられるものとする.

$H_n(x, y; b)$ を状態 $(n; x, y; b)$ のときに最適政策を用いたときに得られる最大期待総利得とする. また $i = 0, 1$ に対して $H_n^i(x, y; b)$ を状態 $(n; x, y; b)$ のときに, まず i を行い, 以後は最適政策を用いたときに得られる最大期待総利得とする. このとき

$$H_n(x, y; b) = \max \{H_n^0(x, y; b), H_n^1(x, y; b)\}$$

ここに

$$H_n^0(x, y; b) = E[(Z \wedge b) + H_{n-1}(\phi(x, y; Z \wedge b, b), \psi(x, y; Z \wedge b, b); b) | x, y; b]$$

および

$$H_n^1(x, y; b) = E[Z \wedge b] + H_{n-1}(x, y; b)$$

となる.

この場合には, 観察値 $z (< b)$ が得られた場合には, 新しい状態は $(n-1; \phi(x, y; z, b), \psi(x, y; z, b); b)$ となる. また観察値 z が b 以上である場合には $(n-1; \phi(x, y; b, b), \psi(x, y; b, b); b)$ となる.

5 3者の比較

5.1 一様分布の場合

実験 a_1 を行った結果は, 区間 $(0, 1)$ 上の一様分布

$$f(z|1) = \begin{cases} 1, & \text{if } 0 < u < 1, \\ 0, & \text{otherwise.} \end{cases}$$

から観察値が得られる. これに対して実験 a_0 を行った結果は, 区間 $(0, u)$ 上の一様分布

$$f(z|u) = \begin{cases} \frac{1}{u}, & \text{if } 0 < z < u, \\ 0, & \text{otherwise.} \end{cases}$$

から観察値が得られる。ここで u の事前分布としてパラメータ (w, α) のパレート分布で与えられるとする。このとき、 u の確率密度関数は

$$g(u|x, y) = \begin{cases} \frac{yx^y}{u^{y+1}}, & \text{if } u > x, \\ 0, & \text{otherwise.} \end{cases}$$

である。また

$$R(x, y) = \frac{1}{2} \frac{yx}{y-1}$$

となる。

区間 $(0, u)$ 上の一様分布の場合には、事前分布のパラメータ x は $x > 1$ となった場合には a_1 は a_0 より劣ることがわかるので、 a_1 を行う必要がなくなる。

観察値が必ず得られる場合には、状態が $(n; x, y)$ で表されているときに、 a_1 を行った結果は $(n-1; x, y)$ となるが、 a_0 を行った結果は確率 1 で $(n-1; \max\{x, z\}, y+1)$ となる。

観察値が得られない可能性がある場合には、状態が $(n; x, y)$ で表されているときに、 a_1 を行った結果は $(n-1; x, y)$ となるが、 a_0 を行った結果は確率 p で $(n-1; \max\{x, z\}, y+1)$ となり、確率 $1-p$ で $(n-1; x, y)$ となる。

5.2 指数分布の場合

実験 $a_i (i = 0, 1)$ の結果が確率密度関数

$$f(z|u_i) = u_i e^{-u_i z} \quad i = 0, 1$$

を持つとする。ただし u_1 は既知であるが、 u_0 は未知で

$$g(u_0|x, y) = \begin{cases} (\Gamma(\alpha))^{-1} x^y u_0^{y-1} e^{-xu_0}, & \text{if } 0 \leq u_0, \\ 0, & \text{otherwise.} \end{cases}$$

を事前分布として持つ。

観察値が必ず得られる場合には、状態が $(n; x, y)$ で表されているときに、 a_1 を行った結果は $(n-1; x, y)$ となるが、 a_0 を行った結果は確率 1 で $(n-1; x+z, y+1)$ となる。

観察値が得られない可能性がある場合には、状態が $(n; x, y)$ で表されているときに、 a_1 を行った結果は $(n-1; x, y)$ となるが、 a_0 を行った結果は確率 p で $(n-1; x+z, y+1)$ となり、確率 $1-p$ で $(n-1; x, y)$ となる。

部分的な情報を利用する場合には、状態が $(n; x, y)$ で表されているときに、 a_1 を行った結果は $(n-1; x, y)$ となるが、 a_0 を行った結果は、もし $0 \leq z \leq b$ ならば $(n-1; x+z, y+1)$ となり、 $b < z$ ならば $(n-1; x+b, y)$ となる。

5.3 ベルヌーイ分布の場合

実験 $a_i (i = 0, 1)$ の結果が $P\{Z = 1\} = 1 - P\{Z = 0\} = u_i$ のとき, これは離散分布である. ここで u_1 は既知であるが, u_0 は未知で事前分布として

$$g(u_0|x, y) = \begin{cases} B(x, y)^{-1} u_0^{x-1} (1 - u_0)^{y-1}, & \text{if } 0 \leq u_0, \\ 0, & \text{otherwise.} \end{cases}$$

を仮定できるものとする. この場合には,

$$\phi(x, y; Z) = \begin{cases} x + 1, & \text{if } Z = 1, \\ x, & \text{if } Z = 0, \end{cases}$$

および

$$\psi(x, y; Z) = \begin{cases} y, & \text{if } Z = 1, \\ y + 1, & \text{if } Z = 0, \end{cases}$$

であり, また

$$h(z|x, y) = \begin{cases} x(x+y)^{-1}, & \text{if } Z = 1, \\ y(x+y)^{-1}, & \text{if } Z = 0, \end{cases}$$

および

$$R(x, y) = \frac{x}{x+y}$$

となる.

この場合には, $z = 0$ か $z = 1$ かのいずれかで, 部分的な情報は存在しないので, 状態が $(n; x, y)$ のときに実験を開始して, その後のある時点において, 結果が得られていなければ, その実験を中断して終了してしまうので, 観察値が必ず得られるか, 観察値が得られないかのいずれかであり, 3節のような解析は可能である. しかしベルヌーイ分布の場合には, 実験を行った結果は成功か失敗かのいずれかであり, 部分的な情報は存在しないと考えられるので, 4節のような解析は不可能である.

参考文献

- [1] Braden, D. J. and Freimer, M. (1991). Informational dynamics of censored observations. *Management Sciences* **37**, 1390-1404.
- [2] Hamada, T. (1978). A uniform two-armed bandit problem: The parameter of one distribution is known. *Journal of Japan Statistical Society* **8**, 29-35.
- [3] Hamada, T. (1989). A two-armed bandit problem with possibility of no information. *Journal of Information & Optimization Sciences* **10**, 281-302.

- [4] Hamada, T. (1995). An exponential two-armed bandit problem with one arm known under batch sampling. *Journal of Japan Statistical Society* **25**, 205-216.
- [5] Lariviere, M. A. and Porteus, E. L. (1999). Stalking information: Bayesian inventory management with unobserved lost sales. *Management Sciences* **45**, 346-363.
- [6] Scarf, H. E. (1959). Bayes solution of the statistical inventory problem. *Ann. Math. Statist.* **30**, 490-508.
- [7] Scarf, H. E. (1960). Some remarks on Bayes solutions to the inventory problem. *Naval Res. Logist. Quarterly* **7**, 591-596.