数理解析研究所講究録
第 842 巻 1993 年 151-156

Two-stage sequential estimation procedures for the uniform distribution

筑波大・数学　赤平　昌文 (Masafumi Akahira)

1．はじめに

区間 $(0,\theta)$ 上の一様分布から逐次的に標本 $X_1, X_2, \ldots$ がとられるときに、高々dの幅をもつ信頼度$1-\alpha$の信頼区間でθを推定する問題において、必要な標本数 Nを求めるための逐次方式が論じられてきた (例えば Graybill and Connell(1964), Cooke(1971))。ここでは2段階逐次方式について Graybill らとは異なる観点から考察する。

2．設定

区間 $(0,\theta)$ 上の一様分布からの第一段階の標本を $X_1,\ldots,X_m$とし、第二段階の標本を $X_{m+1},\ldots,X_{m+N}$とし、$X_{(m)} := \max_{1\le i\le m} X_i$, $X_{(m+N)} := \max_{1\le i\le m+N} X_i$とするとき、$0<\alpha<1$と適当な関数 gについて

$$P_\theta\left\{X_{(m+N)} < \theta \le g\left(X_{(m+N)}\right)\right\} \ge 1-\alpha, \quad {}^\forall\theta>0 \tag{1}$$

となるように $X_{(m)}$に基づいて Nを、すなわち N を $X_{(m)}$の関数 $h\left(X_{(m)}\right)$として求める。

なお Graybill and Connell(1964) では、$Y_N := \max_{m+1\le i\le m+N} X_i$とおいて、$d>0$, $0<\alpha<1$ について

$$P_\theta\left\{|Y_N-\theta|<d\right\} > 1-\alpha$$

となるように $X_{(m)}$ に基づいて N を求める方式を考えている。

さて、ここではまず関数 $g(x)$ を $\mathbf{R}^1$ 上の狭義単調増加関数とし、また $h(x)$ の関数型としては、$(0,\infty)$ 上の階段関数

$$h(x)=\sum_{k=1}^{\infty} k\chi_{(a_{k-1},a_k]}(x)$$

を考えることにする。ただし $\{a_n\}_{n=0,1,\ldots}$ は正数の単調増加列とする。このとき、(1) を満たすような $N=h\left(X_{(m)}\right)$ を求めるためには、点列 $\{a_n\}_{n=0,1,\ldots}$ を求めればよいことが分かる。

3．2段階逐次方式

前節の (1) を考えるために $P_\theta\left\{g\left(X_{(m+N)}\right)<\theta\right\}$ を計算する。まず

(2) $$\begin{aligned}&P_\theta\left\{g\left(X_{(m+N)}\right)<\theta\right\}\\&=\sum_{n=0}^{\infty}P_\theta\left\{h\left(X_{(m)}\right)=n\right\}P_\theta\left\{g\left(X_{(m+n)}\right)<\theta, X_{(m+n)}=X_{(m)}\middle|h\left(X_{(m)}\right)=n\right\}\\&+\sum_{n=1}^{\infty}P_\theta\left\{h\left(X_{(m)}\right)=n\right\}P_\theta\left\{g\left(X_{(m+n)}\right)<\theta, X_{(m+n)}>X_{(m)}\middle|h\left(X_{(m)}\right)=n\right\}\end{aligned}$$

となる。各 $n=1,2,\ldots$ について

(3) $$\begin{aligned}&P_\theta\left\{g\left(X_{(m+n)}\right)<\theta, X_{(m+n)}=X_{(m)}\middle|h\left(X_{(m)}\right)=n\right\}\\&=P\left\{g\left(X_{(m)}\right)<\theta\,\middle|a_{n-1}<X_{(m)}\leq a_n\right\}\\&=P\left\{X_{(m)}<g^{-1}(\theta)\,\middle|a_{n-1}<X_{(m)}\leq a_n\right\}\\&=\begin{cases}0 & (g^{-1}(\theta)\leq a_{n-1}\text{のとき})\\ \dfrac{P_\theta\left\{a_{n-1}<X_{(m)}<g^{-1}(\theta)\right\}}{P_\theta\left\{a_{n-1}<X_{(m)}\leq a_n\right\}} & (a_{n-1}<g^{-1}(\theta)\leq a_n\text{のとき})\\ 1 & (a_n<g^{-1}(\theta)\text{のとき})\end{cases}\end{aligned}$$

となる。各 $n=1,2,\ldots$ について、$a_{n-1}<g^{-1}(\theta)\leq a_n$ のとき

$$P_\theta\left\{a_{n-1}<X_{(m)}<g^{-1}(\theta)\right\}=\left\{\frac{g^{-1}(\theta)}{\theta}\right\}^m-\left(\frac{a_{n-1}}{\theta}\right)^m$$

であるから、(3) より

(4)
$$P_\theta\left\{g\left(X_{(m+n)}\right)<\theta, X_{(m+n)}=X_{(m)} | h\left(X_{(m)}\right)=n\right\}$$
$$=\begin{cases} 0 & (g^{-1}(\theta) \leq a_{n-1}\text{のとき}) \\ \dfrac{\left\{g^{-1}(\theta)/\theta\right\}^m-(a_{n-1}/\theta)^m}{P_\theta\left\{a_{n-1}<X_{(m)} \leq a_n\right\}} & (a_{n-1}<g^{-1}(\theta) \leq a_n\text{のとき}) \\ 1 & (a_n<g^{-1}(\theta)\text{ のとき}) \end{cases}$$

になる。よって、$h\left(X_{(m)}\right)=n$ と $a_{n-1}<X_{(m)} \leq a_n$ は同値であることを用いて、(4) より

(5)
$$\sum_{n=1}^{\infty} P_\theta\left\{h\left(X_{(m)}\right)=n\right\} P_\theta\left\{g\left(X_{(m+n)}\right)<\theta, X_{(m+n)}=X_{(m)} | h\left(X_{(m)}\right)=n\right\}$$
$$=\sum_{n=1}^{\infty}\left[\chi_{\{a_{n-1}<g^{-1}(\theta) \leq a_n\}}(\theta)\left\{\left(\frac{g^{-1}(\theta)}{\theta}\right)^m-\left(\frac{a_{n-1}}{\theta}\right)^m\right\}\right.$$
$$\left.+\chi_{\{a_n<g^{-1}(\theta)\}}(\theta) P_\theta\left\{a_{n-1}<X_{(m)} \leq a_n\right\}\right]$$
$$=\sum_{n=1}^{\infty}\left[\chi_{\{a_{n-1}<g^{-1}(\theta) \leq a_n\}}(\theta)\left\{\left(\frac{g^{-1}(\theta)}{\theta}\right)^m-\left(\frac{a_{n-1}}{\theta}\right)^m\right\}\right.$$
$$\left.+\chi_{\{a_n<g^{-1}(\theta)\}}(\theta)\left\{\left(\frac{a_n}{\theta}\right)^m-\left(\frac{a_{n-1}}{\theta}\right)^m\right\}\right]$$

となる。また $n=0$ のときは

(6)
$$P_\theta\left\{h\left(X_{(m)}\right)=0\right\} P_\theta\left\{g\left(X_{(m)}\right)<\theta \,|h\left(X_{(m)}\right)=0\right\}$$
$$=P_\theta\left\{g\left(X_{(m)}\right)<\theta, h\left(X_{(m)}\right)=0\right\}$$
$$=P_\theta\left\{0<X_{(m)}<g^{-1}(\theta), 0<X_{(m)} \leq a_0\right\}=H_0(\theta) \quad \text{(say)}$$

となる。

次に、各 $n=1,2,\ldots$ について

(7)
$$P_\theta\left\{g\left(X_{(m+n)}\right)<\theta, X_{(m+n)}>X_{(m)} \,| h\left(X_{(m)}\right)=n\right\}$$
$$=P_\theta\left\{X_{(m)}<X_{(m+n)}<g^{-1}(\theta) \,| a_{n-1}<X_{(m)} \leq a_n\right\}$$
$$=\begin{cases} 0 & (g^{-1}(\theta) \leq a_{n-1}\text{のとき}) \\ \dfrac{P_\theta\left\{X_{(m)}<X_{(m+n)}<g^{-1}(\theta), a_{n-1}<X_{(m)} \leq a_n\right\}}{P_\theta\left\{a_{n-1}<X_{(m)} \leq a_n\right\}} & (a_{n-1}<g^{-1}(\theta)\text{ のとき}) \end{cases}$$

になる。このとき $a_{n-1} < g^{-1}(\theta) \le a_n$ について

(8)
$$P_\theta\left\{X_{(m)} < X_{(m+n)} < g^{-1}(\theta), a_{n-1} < X_{(m)} \le a_n\right\}$$
$$= \binom{m+n}{n}\frac{m}{\theta}\int_{a_{n-1}}^{g^{-1}(\theta)}\left(\frac{x_{(m)}}{\theta}\right)^{m-1}\left(\frac{g^{-1}(\theta)}{\theta}-\frac{x_{(m)}}{\theta}\right)^n dx_{(m)}$$

となり、また $a_n < g^{-1}(\theta)$ について

(9)
$$P_\theta\left\{X_{(m)} < X_{(m+n)} < g^{-1}(\theta), a_{n-1} < X_{(m)} \le a_n\right\}$$
$$= \binom{m+n}{n}\frac{m}{\theta}\int_{a_{n-1}}^{a_n}\left(\frac{x_{(m)}}{\theta}\right)^{m-1}\left(\frac{g^{-1}(\theta)}{\theta}-\frac{x_{(m)}}{\theta}\right)^n dx_{(m)}$$

となる。よって、$h\left(X_{(m)}\right) = n$ と $a_{n-1} < X_{(m)} \le a_n$ が同値であることを用いて、(7),(8),(9) から

(10)
$$\sum_{n=1}^{\infty} P_\theta\left\{h\left(X_{(m)}\right)=n\right\} P_\theta\left\{g\left(X_{(m+n)}\right)<\theta, X_{(m+n)} > X_{(m)} | h\left(X_{(m)}\right)=n\right\}$$
$$=\sum_{n=1}^{\infty}\left[\chi_{\{a_{n-1}<g^{-1}(\theta)\le a_n\}}(\theta)\binom{m+n}{n}\frac{m}{\theta}\int_{a_{n-1}}^{g^{-1}(\theta)}\left(\frac{x_{(m)}}{\theta}\right)^{m-1}\left(\frac{g^{-1}(\theta)}{\theta}-\frac{x_{(m)}}{\theta}\right)^n dx_{(m)}\right.$$
$$\left.+\chi_{\{a_n<g^{-1}(\theta)\}}(\theta)\binom{m+n}{n}\frac{m}{\theta}\int_{a_{n-1}}^{a_n}\left(\frac{x_{(m)}}{\theta}\right)^{m-1}\left(\frac{g^{-1}(\theta)}{\theta}-\frac{x_{(m)}}{\theta}\right)^n dx_{(m)}\right]$$

になる。ゆえに (1),(2),(5),(6),(10) より

(11)
$$P_\theta\left\{g\left(X_{(m+N)}\right)<\theta\right\}$$
$$= H_0(\theta)+\sum_{n=1}^{\infty}\left[\chi_{\{a_{n-1}<g^{-1}(\theta)\le a_n\}}(\theta)\left\{\left(\frac{g^{-1}(\theta)}{\theta}\right)^m-\left(\frac{a_{n-1}}{\theta}\right)^m\right.\right.$$
$$\left.+\binom{m+n}{n}\frac{m}{\theta}\int_{a_{n-1}}^{g^{-1}(\theta)}\left(\frac{x_{(m)}}{\theta}\right)^{m-1}\left(\frac{g^{-1}(\theta)}{\theta}-\frac{x_{(m)}}{\theta}\right)^n dx_{(m)}\right\}$$
$$+\chi_{\{a_n<g^{-1}(\theta)\}}(\theta)\left\{\left(\frac{a_n}{\theta}\right)^m-\left(\frac{a_{n-1}}{\theta}\right)^m\right.$$
$$\left.\left.+\binom{m+n}{n}\frac{m}{\theta}\int_{a_{n-1}}^{a_n}\left(\frac{x_{(m)}}{\theta}\right)^{m-1}\left(\frac{g^{-1}(\theta)}{\theta}-\frac{x_{(m)}}{\theta}\right)^n dx_{(m)}\right\}\right]$$
$$\le \alpha, \qquad {}^\forall\theta > 0$$

となるように $\{a_n\}_{n=0,1,\ldots}$ を定めればよい。

特に、$d>0$ を固定して、$g(x)=x+d$ とすると $g^{-1}(\theta)=\theta-d$ となるから、(11) より

$$\begin{aligned}
&P_\theta\left\{X_{(m+N)}+d<\theta\right\}\\
&=H_0(\theta)+\sum_{n=1}^{\infty}\left[\chi_{[a_{n-1}+d,a_n+d)}(\theta)\left\{\left(\frac{\theta-d}{\theta}\right)^m-\left(\frac{a_{n-1}}{\theta}\right)^m\right.\right.\\
&\quad\left.+\binom{m+n}{n}\frac{m}{\theta}\int_{a_{n-1}}^{\theta-d}\left(\frac{x_{(m)}}{\theta}\right)^{m-1}\left(\frac{\theta-d}{\theta}-\frac{x_{(m)}}{\theta}\right)^n dx_{(m)}\right\}\\
&\quad+\chi_{[a_n+d,\infty)}(\theta)\left\{\left(\frac{a_n}{\theta}\right)^m-\left(\frac{a_{n-1}}{\theta}\right)^m\right.\\
&\quad\left.\left.+\binom{m+n}{n}\frac{m}{\theta}\int_{a_{n-1}}^{a_n}\left(\frac{x_{(m)}}{\theta}\right)^{m-1}\left(\frac{\theta-d}{\theta}-\frac{x_{(m)}}{\theta}\right)^n dx_{(m)}\right\}\right]\\
&=H_0(\theta)+\sum_{n=1}^{\infty}\left[\chi_{[a_{n-1}+d,a_n+d)}(\theta)\left\{A_n(\theta)\right\}+\chi_{[a_n+d,\infty)}(\theta)\left\{B_n(\theta)\right\}\right]\quad\text{(say)}
\end{aligned}$$

また、このとき

$$H_0(\theta)=\chi_{[d,a_0+d)}(\theta)\left(\frac{\theta-d}{\theta}\right)^m+\chi_{[a_0+d,\infty)}(\theta)\left(\frac{a_0}{\theta}\right)^m$$

になる。そこで

$$P_\theta\left\{X_{(m+N)}+d<\theta\right\}<\alpha,\qquad {}^\forall\theta>0$$

となる $\{a_n\}_{n=0,1,\ldots}$ を、例えば、まず $\left(\dfrac{a_0}{a_0+d}\right)^m=\dfrac{\alpha}{2}$ となる a_0 をとり、以後、逐次的に

$$\begin{aligned}
&\left(\frac{a_0}{a_n+d}\right)^m+B_1(a_n+d)+B_2(a_n+d)+\cdots+B_n(a_n+d)\\
&=\frac{\alpha}{2}+\frac{\alpha}{2^2}+\cdots+\frac{\alpha}{2^{n+1}}
\end{aligned}$$

となる最小の a_n をとればよい。

参考文献

Cooke, P. J. (1971). Sequential estimation in the uniform density. *J. Amer. Statist. Assoc.*, **66**, 614-617.

Graybill, F. A. and Connell, T. L. (1964). Sample size required to estimate the parameter in the uniform density within d units of the true value. *J. Amer. Statist. Assoc.*, **59**, 550-556.