No. 627

一様分布Ｕ（θ, $\theta+1$)の統計的推測ＩＩ
（Ｄ．Ｐ．Ｎｏ.507の改定、及び、追加補足版)

by

野上佳子

May 1995

一様分布Ｕ［θ，θ＋１）の統計的推測ＩＩ

（Ｄ．Ｐ．　Ｎｏ．５０７の改定、及び、追加補足版）

野　上　佳　子

概要.

Ｄ．Ｐ．　Ｎｏ．５０７　では、ジャックナイフ法による推定量Ｊ（$\hat{\theta}$）（ｃｆ．（６））について、統計的推測を、行ったが、不偏な推定量Ｙ＝（$X_{(1)}$ ＋$X_{(n)}$ －１）／２　については、分布の導出に誤りがあった。この論文では、その誤りを正し、上と同じ推測をし、また、両方の推定量が、同じ極限分布を持つことを、確かめている。

§0. 初めに.

まず、集合Aに対して、指示関数$I_A(x)$を、$I_A(x)=1$ $(x \varepsilon A)$; $=0$ $(x \notin A)$と定義する。密度関数$f(x \mid \theta)=I_{[\theta,\theta+1)}(x)$をもつ母集団から、大きさnのランダム標本$X_1, \ldots, X_n$をとり、これを、大きさの順に並べたものを、$X_{(1)} \leqq X_{(2)} \leqq \ldots \leqq X_{(n)}$とする。帰無仮説$H_0 : \theta=\theta_0$と 対立仮説$H_1$: $\theta \neq \theta_0$を検定するために、不偏推定量$Y=(X_{(1)}+X_{(n)}-1)/2$を用いて、基軸法により、棄却域Cを作り、検定力関数を計算する。これを、D. P. No. 507で、偏りのある最尤推定量(m. l. e.) $\hat{\theta}=X_{(1)}$ $(X_{(n)}-1$でもよい)にジャックナイフ法を施して、バイアスを修正した推定量$J(\hat{\theta})$を用いて作った棄却域C_Jに基ずく検定力関数と、比較する。

結果的には、Yに基ずく検定の方が、検定力が高いが、いずれも、漸近的に、一様最強力不偏検定となり、$n \to \infty$とした時のそれぞれの漸近分布が、同じになることから、ジャックナイフ法がかなり良い手法であることを示すことになってしまったが、それはそれとして、面白い結果と思い、改めて、整理することにした。(D. P. No. 507も参照されたい。) 尚、$J(\hat{\theta})$の漸近分布についても訂正して、第4節に付け加えた。

また、$\doteq$は、右辺を、左辺で定義することを示す。

§1. 信頼区間.

αを、$0<\alpha<1$なる実数とする。 この節では、$Y=(X_{(1)}+X_{(n)}-1)/2$を用いて、θの最も短い(信頼係数$1-\alpha$の)信頼区間を作る。

$(X_{(1)}, X_{(n)})$の同時密度から、変数変換法により、Yの密度関数$g(y \mid \theta)$は、容易に、次のように求まる。

(1) $$g(y \mid \theta)=n(1-2|y-\theta|)^{n-1} \quad (-1/2<y-\theta<1/2).$$

$Q \doteq Y-\theta$は、基軸量なので、基軸法(A. M. Mood, F. A. Graybill, & D. C. Bose (1974))により、信頼係数$1-\alpha$のθの信頼区間を求めると、

(2) $$(Y-r, Y+r) \qquad (ここで、r=(1-\alpha^{1/n})/2。)$$

が得られる。 この区間の長さは、$1-\alpha^{1-n}$ で、$J(\hat{\theta})$（cf.（6））を用いた信頼区間の長さ $(2-n^{-1})(1-\alpha^{1/n})$ と比較すると、かなり短い。

§2. 検定力関数.

この節では、$H_0:\theta=\theta_0$ と、$H_1:\theta\neq\theta_0$（θ_0：実数）を検定するための棄却域Cと、検定力関数を求める。 （2）式から、Cを、

$$C=\{y:\ y\leqq\theta_0-r \quad \text{或いは} \quad y\geqq\theta_0+r\}$$

なる棄却域とする。 この棄却域をもつ検定の検定力関数を、$\pi_C(\theta)=P_\theta(C)$ とすると、$\pi_C(\theta)$ は、容易に、下のように求まる。 （rは、（2）式を、参照。）

$$(3)\qquad \pi_C(\theta)=\begin{cases} 1, & (\theta<\theta_0-r-1/2), \\ 1-\{1-2(\theta_0-r-\theta)\}^n/2, & (\theta_0-r-1/2\leqq\theta<\theta_0+r-1/2) \\ 1-[\{1-2(\theta_0-r-\theta)\}^n-\{1-2(\theta_0+r-\theta)\}^n]/2, & (\theta_0+r-1/2\leqq\theta<\theta_0-r) \\ [\{1-2(\theta-\theta_0+r)\}^n+\{1-2(\theta_0+r-\theta)\}^n]/2, & (\theta_0-r\leqq\theta<\theta_0+r), \\ 1-[\{1+2(\theta_0+r-\theta)\}^n-\{1-2(\theta-\theta_0+r)\}^n]/2, & (\theta_0+r\leqq\theta<\theta_0-r+1/2) \\ 1-\{1+2(\theta_0+r-\theta)\}^n/2, & (\theta_0-r+1/2\leqq\theta<\theta_0+r+1/2), \\ 1, & (\theta_0+r+1/2\leqq\theta). \end{cases}$$

（ここで、D. P. No. 507 のTable 2.1 の$\alpha^{1/n}$ の値をみるとわかるように、$\alpha\geqq.001$, $n\geqq10$ に対しては、$\alpha^{1/n}\geqq1/2$である。従って、（3）の計算は、$\alpha^{1/n}\geqq1/2$に基づいている。）

$\theta<\theta_0$ なる θ に対しては、$\pi'_c(\theta)<0$、$\theta>\theta_0$ なる θ に対しては、$\pi'_c(\theta)>0$、かつ、$\pi'_c(\theta_0)=0$、また、$\pi_c(\theta_0)=\alpha$、$\pi_c(+\infty)=\pi_c(-\infty)=1$ であることは、簡単に示せるので、$\pi_c(\theta)$ は、θ の、下に凸な関数であり、棄却域Cに基づく検定は、大きさ α の不偏検定といえる。
また、$n\to\infty$ の時、$\theta \neq \theta_0$ なる全ての θ に対して $\pi_c(\theta)\uparrow 1$ なので、漸近的に、一様最強力（UMP）検定である。

次に、Yの漸近分布を求める。

§3.　漸近分布.

この節では、$Z_n=(Y-E(Y))/\sqrt{Var(Y)}$ の漸近分布を、求める。
（1）式より、容易に、$E(Y)=\theta$ と $Var(Y)=(2(n+1)(n+2))^{-1}$ が求まる。（1）式から、Z_n の密度関数 $g_{Z_n}(z)$ が、次のように求まる。（但し、$c_0=\sqrt{2(n+1)(n+2)}$）

$$(4)\qquad g_{Z_n}(z)=(n/c_0)(1-(2|z|/c_0))^{n-1}\times I_{(-1/2,1/2)}(z/c_0)$$

従って、

$$(5)\qquad \lim_{n\to\infty} g_{Z_n}(z) = 2^{-1/2}e^{-\sqrt{2}|z|}I_{(-\infty,\infty)}(z)\quad(=\Psi(z))$$

となる。

次の節では、$J(\hat{\theta})$ の漸近分布（D. P.　No. 507 のp. 8における $h^*(z|\theta)$ は、誤り。）を求める

§4.　$J(\hat{\theta})$ の漸近分布.

初めに、Quenouille（1949, 1956）の方法を用いて $J(\hat{\theta})$ を導出する。θ を、$\hat{\theta}=t(X_1,\cdots,X_n)=X_{(1)}$ で推定すると、

$$\hat{\theta}^{i} \doteq t(X_1, \cdots, X_{i-1}, X_{i+1}, \cdots, X_n)$$

$$\doteq \begin{cases} X_{(1)}, & X_i \neq X_{(1)} \text{ の時} \\ X_{(2)}, & X_i = X_{(1)} \text{ の時、} \quad (i = 1, 2, \cdots, n). \end{cases}$$

$\overline{\hat{\theta}^{i}} = n^{-1}\Sigma_{i=1}^{n}\hat{\theta}^{i}$ と定義すると、θのジャックナイフ推定量は、

(6) $$J(\hat{\theta}) = n\hat{\theta} - (n-1)\overline{\hat{\theta}^{i}}$$

となる。簡単のために、$W \doteq J(\hat{\theta})$ とおく。

ここで、$U_n = (W - E_\theta(W)) / \sqrt{Var_\theta(W)}$ の漸近分布を求める。Wの密度関数$h(w \mid \theta)$は、

(7) $$h(w \mid \theta) = n(ab^{n-1})^{-1}(w - \theta + b)^{n-1} I_{(\theta-b,\theta)}(w)$$
$$+ na^{-1}(1 + \theta - w)^{n-1} I_{[\theta,\theta+1)}(w)$$

と表わされる。(ここで、$a \doteq 2 - n^{-1}$ と $b \doteq 1 - n^{-1}$ である。) これより、$E_\theta(W)$と$Var_\theta(W)$は、

(8) $$E_\theta(W) = \theta + (n(n-1))^{-1},$$

(9) $$V_n \doteq Var_\theta(W) = (-4\theta)/\{n(n^2-1)\}$$
$$+ (2n^3 - 6n^2 + 7n - 9)/\{n(n+1)(n+2)(n-1)^2\}$$

となる。 また、(7)式から、変数変換法により、U_nの密度関数$h_{U_n}(u)$は、次のように、求まる。

(10) $$h_{U_n}(u) = na^{-1}\{1 - u\sqrt{V_n} - 1/(n(n-1))\}^{n-1}\sqrt{V_n} \times I_{(-1/(n(n-1)),\,1-1/(n(n-1)))}(u\sqrt{V_n})$$
$$+ n(ab^{n-1})^{-1}\{b + u\sqrt{V_n} + 1/(n(n-1))\}^{n-1}\sqrt{V_n} \times I_{(-(b+1/(n(n-1))),\,-1/(n(n-1)))}(u\sqrt{V_n})$$

従って、

(11)　$\lim_{n\to\infty} h_{U_n}(u) = 2^{-1/2} e^{-\sqrt{2}|z|} I_{(-\infty,\infty)}(u) \quad (=\Psi(u))$

これは、3節の結果と同じである。

§5.　結論.

D. P. No. 507 p. 4 の(2. 1)式の$\pi_{cJ}(\theta)$と、本論文の(3)式の$\pi_c(\theta)$を比較すると、$\pi_{cJ}(\theta)<1$のθの範囲は、$\pi_c(\theta)<1$のθの範囲$2-\alpha^{1/n}$の$2-n^{-1}$倍も広く、Cを用いた検定の方が、検定力が高いといえそうである。これを調べるために、付表に、2つの検定力関数の差$(A \doteq)\ \pi_c(\theta)-\pi_{cJ}(\theta)$の値を、各$\alpha=.10$, $.05$, $.01$と各$n=10$, 50, 100について計算してみた。(表の値は、左端のθの各点におけるAの値である。)この表からは、$\pi_c(\theta)$が、$\pi_{cJ}(\theta)$より、一様に検定力が、高いことが、推察される。

にもかかわらず、$J(\hat{\theta})$についての解析を紹介するのは、もとの推定量$\hat{\theta}=X_{(1)}$の分布は、極端に歪んだ分布でありながら、ジャックナイフ法により、修正することによって、棄却域Cを持つ両側検定にかなり近付けるという点は、注目に値する。
また、Z_nとU_nが共に同じ極限分布$\Psi(\cdot)$をもつ点も、興味深いと思う。

現在、上の結果を拡張中である。

参考文献：

1) Mood, A. M., Graybill, F. A. &Bose, D. C. (1974) Introduction to the Theory of Statistics, McGraw-Hill Kogakusha, LTD., 3rd ed.

2) Nogami, Y. (1992). "A Statistical Inference under the Uniform Distribution U[θ, θ+1)." Inst. of Socio-Economic Planning, U. of Tsukuba, D. P. Series No. 507, December pp. 1−9

3) Quenouille, M. (1949). "Approximate tests of correlation in time series." J. Roy. Statist. Soc. Ser. B, 11, pp. 68−84

4) ----- (1956). "Note on bias in estimation" Biometrika 43, pp. 353−360.

付表.

$A = \pi_C - \pi_{CJ}$ の表

θ \ n \ α	α=.10			α=.05			α=.01		
	n=10	50	100	n=10	50	100	n=10	50	100
$\theta_0 - 1 + \alpha^{1/n}/2$	+	+	+	+	+	+	+	+	+
$\theta_0 - \alpha^{1/n}/2$	.0524	.4456 −14	.1295 −29	.1213	.1591 −13	.2074 −28	.6045	.6031 −12	.3083 −27
$\theta_0 - r$	.3124	.3491	.3006	.3578	.3849	.3845	.4213	.4466	.4318
θ_0	0	0	0	0	0	0	0	0	0
$\theta_0 + r$	.3195	.3504	.3011	.3669	.3864	.3853	.4319	.4491	.4325
$\theta_0 + \alpha^{1/n}/2$	.0473	.2227 −14	.5414 −30	.1279	.8453 −14	.9303 −29	.8326	.3776 −12	.1478 −27
$\theta_0 + 1 - \alpha^{1/n}/2$	+	+	+	+	+	+	+	+	+