棄却検定の比較表 (2)

石川栄助

A Comparative Table to Make Sample Criteria for Testing Rejection (2)

EISUKE ISHIKAWA

§ 0. はしがき
§ 1. SMIRNOV の棄却限界値
§ 2. SMIRNOV の限界値と GRUBBS の限界値
§ 3. SMIRNOV の限界値と CHAMOVA の限界値
§ 4. SMIRNOV の限界値と THOMPSON の限界値

§ 0. はしがき

規準正規母集団 $N(0,1)$[1] より大いさ n の任意の標本を $X_1, X_2, \cdots\cdots, X_n$ とし，次の統計量：

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad , \quad X' = \frac{1}{n-1}\sum_{i=1}^{n-1} X_i \qquad \cdots\cdots (0.1)$$

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2, \qquad \tau = \frac{X_n-\bar{X}}{s} \qquad \cdots\cdots (0.2)$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2, \qquad \tau' = \frac{X_n-\bar{X}}{s} \qquad \cdots\cdots (0.3)$$

を求めるとき

(1) X_n をこの集団から離れているとし，

$$P_r(|\tau| \geqq \tau_t) = \frac{2\alpha}{n} \qquad \cdots\cdots (0.4)$$

なる棄却限界値 τ_t を求めたのが THOMPSON の値である[2].

(2) X_n を最大（または最小）とし，順序統計量として

$$P_r(\tau \geqq \tau_s) = \alpha \qquad \cdots\cdots (0.5)$$

として棄却限界値 τ_s を求めたのが，SMIRNOV-GRUBBS の値である[3][4].

(3) CHAMOVA は (2) の τ の代り τ' を用いて，

$$P_r(\tau' \geqq \tau'_c) = \alpha \qquad \cdots\cdots (0.6)$$

なる限界値 τ'_c を求めている[3].

同じ有意水準 α に対して，τ_t, τ_s, τ'_c の間に大体 $\tau_t \fallingdotseq \tau_s$ が成立する．ここに二つの限界値を比較し，上の関係を具体的に示し，更に

変域：

$$\sqrt{\frac{n-2}{2}} < \tau_0 < \sqrt{n-1}$$

1) 一般正規母集団 $N(m,\sigma^2)$ としてもよい．

2) 参考文献 (3)．，3)，4) 参考文献 (1)，(2)

に於て両式が完全に一致する事を証したい.

§1. SMIRNOV の棄却限界値

母集団 N (0, 1) からの標本：

$$X_1 \leq X_2 \leq \cdots\cdots\cdots\cdots \leq X_{n-1} \leq X_n$$

に於て，統計量 τ を求め

$$\alpha = P_r(\tau \geqq \tau_s) = 1 - Q_n(\tau_s) \qquad \cdots\cdots\cdots\cdots (1.1)$$

なる τ_s を求めたのが SMIRNOV である.

ここに

$$Q_n(\tau_s) = \frac{n!}{(\sqrt{2\pi})^n} \int_{D_n} e^{-\frac{1}{2}\Sigma X_i^2} dX_1 dX_2 \cdots\cdots dX_n \qquad \cdots\cdots\cdots\cdots (1.2)$$

但し　D_n： $-\infty < X_1 \leqq X_2 \leqq \cdots\cdots\cdots \leqq X_n \leqq \bar{X} + s\tau_s$

今次の直交変換を行い

$$\left.\begin{aligned}
X_1 &= \bar{X} - \frac{1}{\sqrt{1.2}} y_1 - \frac{1}{\sqrt{2.3}} y_2 - \cdots\cdots\cdots - \frac{1}{\sqrt{(n-1)n}} y_{n-1} \\
X_2 &= \bar{X} + \frac{1}{\sqrt{1.2}} y_1 - \frac{1}{\sqrt{2.3}} y_2 - \cdots\cdots\cdots - \frac{1}{\sqrt{(n-1)n}} y_{n-1} \\
X_3 &= \bar{X} \qquad\qquad\quad + \frac{1}{\sqrt{2.3}} y_2 - \cdots\cdots\cdots - \frac{1}{\sqrt{(n-1)n}} y_{n-1} \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
X_n &= \bar{X} \qquad\qquad\qquad\qquad\qquad\qquad + \frac{1}{\sqrt{(n-1)n}} y_{n-1}
\end{aligned}\right\} \cdots\cdots (1.3)$$

とおき，更に極変換：

$$\left.\begin{aligned}
y_1 &= \sqrt{n}\, s \cos\varphi_{n-2} \cos\varphi_{n-3} \cdots\cdots\cdots \cos\varphi_1 \\
y_2 &= \sqrt{n}\, s \cos\varphi_{n-2} \cos\varphi_{n-3} \cdots\cdots\cdots \sin\varphi_1 \\
&\cdots\cdots\cdots\cdots\cdots \\
y_{n-1} &= \sqrt{n}\, s \sin\varphi_{n-2}
\end{aligned}\right\} \cdots\cdots (1.4)$$

とおけば

$$Q_n(\tau_s) = \frac{n!\,\Gamma\left(\frac{n-1}{2}\right)}{2(\sqrt{\pi})^{n-1}} \int_{D_{n-2}} \prod_{i=2}^{n-2} \cos^{i-1}\varphi_i \, d\varphi_1 \, d\varphi_2 \cdots\cdots\cdots d\varphi_{n-2}$$

ここに

$$\sin\varphi_k = \frac{t_k}{\sqrt{k+1}} \qquad (k = 1, 2, \cdots\cdots, n-2)$$

この積分を行い

$$\left.\begin{aligned} Q_n(\tau_s) &= c_n \int_{\frac{1}{\sqrt{n-1}}}^{\tau_s} \left(1-\frac{t^2_{n-2}}{n-1}\right)^{\frac{n-4}{2}} dt_{n-2} \cdot \prod_{k=2}^{n-2} \int_{\frac{1}{\sqrt{k}}}^{a_k} \left(1-\frac{t^2_{k-1}}{k}\right)^{\frac{k-3}{2}} dt_{k-1} \\ \text{ここに}\quad c_n &= \frac{n\,!\,\Gamma\left(\frac{n-1}{2}\right)}{2\left(\sqrt{\pi}\right)^{n-1}} \prod_{k=1}^{n-2} \frac{1}{\sqrt{k+1}} \end{aligned}\right\} \cdots\cdots (1.5)$$

$$a_k = a_k(\mathbf{t}_k) = min \left\{ \sqrt{k+1},\ \sqrt{\frac{k+2}{k+1-t_k^2}}\, t_k \right\}$$

簡単のために

$$\varphi_n(\tau) = \frac{\Gamma\left(\frac{n-1}{2}\right)}{\sqrt{(n-1)\pi}\ \Gamma\left(\frac{n-2}{2}\right)} \left(1-\frac{\tau^2}{n-1}\right)^{\frac{n-4}{2}} \qquad \cdots\cdots\cdots\cdots\cdots\cdots (1.6)$$

とおけば上式は

$$\left.\begin{aligned} Q'_n(\tau_s) &= n\,\varphi_n(\tau_s)\,Q_{n-1}\left\{\delta(\tau_s)\right\} \\ \delta(\tau_s) &= \sqrt{\frac{n\,\tau_s^2}{n-1-\tau_s^2}} \end{aligned}\right\} \qquad \cdots\cdots\cdots\cdots\cdots\cdots (1.7)$$

となる．ただし $\dfrac{1}{\sqrt{n-1}} < \tau_s < \sqrt{n-1}$

上の積分は区間 $\left(\sqrt{\dfrac{1}{n-1}},\ \sqrt{\dfrac{2}{n-2}}\right)$, $\left(\sqrt{\dfrac{2}{n-2}},\ \sqrt{\dfrac{3}{n-3}}\right)$, …………,

$\left(\sqrt{\dfrac{n-2}{2}},\ \sqrt{n-1}\right)$ によつて積分しなければならない．

§ 2. SMIRNOV と GRUBBS の棄却限界値

GRUBBS は次の統計量：

$$\left.\begin{aligned} \bar{\mathbf{x}}' &= \frac{1}{n-1}\sum_{i=1}^{n-1} X_i\ , \qquad S' = \sum_{i=1}^{n-1}\left(X_i - \bar{\mathbf{x}}'\right)^2 \\ \bar{\mathbf{x}} &= \frac{1}{n}\sum_{i=1}^{n} X_i\ , \qquad S = \sum_{i=1}^{n}\left(X_i - \bar{\mathbf{x}}\right)^2 \end{aligned}\right\} \qquad \cdots\cdots\cdots\cdots\cdots\cdots (2.1)$$

を求めた．しかる時，簡単に次の関係式をうる．

$$\sqrt{(n-1)\left(1-\frac{S'}{S}\right)} = \tau \qquad \cdots\cdots\cdots\cdots\cdots\cdots (2.2)$$

GRUBBS は S'/S の分布を研究し，この値の α-Point の値を求め，更に上式に代入して，τ の限界値を独立に定めた．この値が SMIRNOV の限界値 τ_s に一致することは § 3. によつて明かである．第1表の α_s に応ずる τ は GRUBBS の求めた限界値である．

第 1 表　SMIRNOV-GRUBBS の τ_s と THOMPSON の τ_t

α \ n	0.05		0.025		0.1	
	$\alpha_s=0.05$	$\alpha_t=\frac{0.10}{n}$	$\alpha_s=0.02$	$\alpha_t=\frac{0.05}{n}$	$\alpha_s=0.1$	$\alpha_t=\frac{0.2}{n}$
3	1.41	1.41	1.41	1.41	1.41	1.41
4	1.69	1.69	1.71	1.71	1.65	1.65
5	1.87	1.87	1.92	1.92	1.79	1.79
6	2.00	2.00	2.07	2.07	1.89	1.90 (*)
7	2.09	2.09	2.18	2.18	1.97	1.97
8	2.17	2.17	2.27	2.27	2.04	2.04
9	2.24	2.24	2.35	2.35	2.10	2.10
10	2.29	2.30 (*)	2.41	2.41	2.15	2.14 (*)
11	2.34	2.34	2.47	2.47	2.19	2.19
12	2.39	2.39	2.52	2.52	2.23	2.23
13	2.43	2.43	2.56	2.57 (*)	2.26	2.26
14	2.46	2.46	2.60	2.60	2.30	2.30
15	2.49	2.50 (*)	2.64	2.64	2.33	2.33
16	2.52	2.52	2.67	2.67	2.35	2.36 (*)
17	2.55	2.55	2.70	2.70	2.38	2.38
18	2.58	2.58	2.73	2.73	3.40	2.40
19	2.60	2.60	2.75	2.76 (*)	2.43	2.43
20	2.62	2.63 (*)	2.78	2.78	2.45	2.45
21	2.64	2.65 (*)	2.80	2.80	2.47	2.47
22	2.66	2.66	2.82	2.82	2.49	2.49

1)　* 印は 0.01 の差がある.

2)　表の値は $P_r(\tau \geqq \tau_0)=\alpha_s$,　$P_r(\tau \geqq \tau_0)=\alpha_t$ の τ_0 である.
これを夫々 τ_s, τ_t とおいた.

§ 3.　SMIRNOV の限界値と CHAMOVA の限界値

CHAMOVA は SMIRNOV の理論によつて統計量：(0.3) 式の τ' を用いて，棄却限界値 τ'_c を $\alpha=0.1$,　0.075,　0.05　0.025　について求めた，比較のために次に掲げる.

第 2 表　CHAMOVA の τ'_c

n \ α	0.1	0.05	0.025
3	1.15	1.15	1.15
4	1.42	1.46	1.48
5	1.60	1.67	1.72
6	1.73	1.82	1.89
7	1.83	1.96	2.02
8	1.91	2.03	2.13
9	1.98	2.11	2.21
10	2.03	2.18	2.29
11	2.09	2.23	2.36
12	2.13	2.29	2.41
13	2.17	2.33	2.47
14	2.21	2.37	2.50
15	2.25	2.41	2.55
16	2.28	2.44	2.58
17	2.31	2.48	2.62
18	2.34	2.50	2.66
19	2.36	2.53	2.68
20	2.38	2.56	2.71

しかして，(0.2) 式 (0.3) 式より

$$\frac{\tau'_c}{\tau_s} = \frac{\mathrm{s}}{s} = \sqrt{\frac{n-1}{n}} \qquad \cdots\cdots (3.1)$$

$$\therefore \quad \tau'_c = \sqrt{\frac{n}{n-1}}\ \tau_s \qquad \cdots\cdots (3.2)$$

$$\tau_s = \sqrt{\frac{n}{n-1}}\ \tau'_c \qquad \cdots\cdots (3.3)$$

をうる．即ち SMIRNOV の限界値と CHAMOVA の限界値の間に (3.1) の関係が成立し，一方から容易に他方が得られる．これから SMIRNOV の限界値を得て，GRUBBS の限界値に一致することがわかる．

§ 4. SMIRNOV-GRUBBS の棄却限界値と THOMPSON の棄却限界値

第 1 表の α_t に応ずる τ_t は次の式によつた THOMPSON の棄却限界値である．この場合の有意水準 α_t は $2\alpha/n$ になつている．

$$P_r\left(|\tau|\geqq\tau_t\right)=\frac{2\Gamma\left(\frac{n-1}{2}\right)}{\sqrt{\pi(n-2)}\ \Gamma\left(\frac{n-2}{2}\right)}\int_{t_o}^{\infty}\left(1+\frac{t^2}{n-2}\right)^{-\frac{n-1}{2}}dt=\frac{2\alpha}{n}\qquad\cdots\cdots\cdots(4.1)$$

但し　$t_o=\sqrt{\dfrac{(n-2)\,\tau_t^2}{n-1-\tau_t^2}}$

これに対する種々の批判がなされてあるが，第一表を見ればわかる様に，αに対するτ_sと$2\alpha/n$に対するτ_tとは殆んど相等しい．
即ち

$$\alpha_t\fallingdotseq\frac{2}{n}\,\alpha_s\qquad\cdots\cdots\cdots\cdots\cdots\cdots(4.2)$$

この関係が $\dfrac{1}{\sqrt{n-1}}<\tau_o<\sqrt{n-1}$ なる全域で成立するかどうかは完証出来なかつたが，

$\sqrt{\dfrac{n-2}{2}}<\tau_0<\sqrt{n-1}$ の変域で，等式の成立すことは次のように証明出来た．

(1.7) 式より

$$\alpha_s=P_r\left(\tau\geqq\tau_o\right)=\frac{n\ \Gamma\left(\frac{n-1}{2}\right)}{\sqrt{(n-1)\pi}\ \Gamma\left(\frac{n-2}{2}\right)}\int_{\tau_o}^{\sqrt{n-1}}\left(1-\frac{\tau^2}{n-1}\right)^{\frac{n-4}{2}}d\tau$$

今　$t_o^2=\dfrac{(n-2)\,\tau^2}{n-1-\tau^2}$　　とおけば

$$\alpha_s=\frac{n\,\Gamma\left(\frac{n-1}{2}\right)}{\sqrt{(n-1)\pi}\ \Gamma\left(\frac{n-2}{2}\right)}\int_{t_o}^{\infty}\left(1+\frac{t^2}{n-2}\right)^{-\frac{n-1}{2}}\sqrt{\frac{n-1}{n-2}}\,dt$$

$$=\frac{n\ \Gamma\left(\frac{n-1}{2}\right)}{\sqrt{(n-2)\pi}\ \Gamma\left(\frac{n-2}{2}\right)}\int_{t_o}^{\infty}\left(1+\frac{t^2}{n-2}\right)^{-\frac{n-1}{2}}dt$$

(4.1) 式と比較して

$$\alpha_s=\frac{n}{2}\,\alpha_t$$

となる．従つて $\alpha=0.1$, 0.05, 0.025 に対し $n=11$, 14, 16 までは完全に一致することになる．よつて実用的には SMIRNOV の τ_s の代用として THOMPSON の τ_t を用いることが出来る．（未完）

参　考　文　献

(1) FRANK. E. GRUBBS : Sample Criteria for testing outlying observations; Ann. of Math. Stat. vol 21 (1950).

(2) SMIRNOV : On the Estimation of the maximum term in a series of observations; Doklady (1941).

(3) 統計科学研究会編：統計数値表 I (1943)
(4) ―――――――― 新編統計数値表 (1952)
(5) 中山伊知郎編：統計学辞典 (1957)
(6) 石川栄助：棄却検定の比較表，岩手大学学芸学部研究年報　第9巻 (1955)
(7) 小川潤次郎：近代数理統計学序説 (1954)

Summary

For testing the significance of the largest (or smallest) observation in a sample of size n from $N(0,1)$, we propose the statistic

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad \bar{X}' = \frac{1}{n-1}\sum_{i=1}^{n-1} X_i$$

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2 \qquad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

$$S = \sum_{i=1}^{n}(X_i-\bar{X}^2) \qquad S' = \sum_{i=1}^{n-1}(X_i-\bar{X}')^2$$

$$\tau = \frac{X_n-\bar{X}}{s} \qquad \tau' = \frac{X_n-\bar{X}}{s}$$

where $X_1 \leqq X_2 \leqq \cdots\cdots \leqq X_n$

Here these manners may be stated as follows :

THOMPSON : $\alpha_t = P_r\,(|\tau| \geqq \tau_o)$ or $P_r\,(|\tau| \geqq \tau_t) = \dfrac{2\alpha}{n}$

SMIRNOV : $\alpha_s = P_r\,(\tau \geqq \tau_o)$ or $P_r\,(\tau \geqq \tau_s) = \alpha$

GRUBBS : $\alpha_s = \sqrt{(n-1)\left(1-\dfrac{S'}{S}\right)}$

CHAMOVA : $P_r\,(\tau' \geqq \tau'_c) = \alpha$

Concerning τ_s and τ'_c we have relation as follows :

$$\frac{\tau'_c}{\tau_s} = \sqrt{\frac{n-1}{n}}$$

And for the interval $\sqrt{\dfrac{n-2}{2}} < \tau_o < \sqrt{n-1}$ we shall have

$$\alpha_t = \frac{2}{n}\alpha_s$$