

6

U統計量の分布関数の収束速度について

静大 工学部 多賀 保志

§ 1. 序

$X_1, X_2, \dots, X_n$ は互いに独立で同一分布 $F(x)$ にしたがう確率変数で, $F(x)$ は 3 次の絶対積率 β_3 , 分散 σ^2 , 平均 0 をもつとすれば, $Y_n = (\sigma\sqrt{n})^{-1} \sum_{i=1}^n X_i$ の分布関数 $G_n(y)$ は標準正規分布 $\Phi(y)$ に一様に収束し, かつすべての n について

$$(1.1) \quad \sup_{-\infty < y < \infty} |G_n(y) - \Phi(y)| < \frac{C \beta_3}{\sqrt{n} \sigma^3}, \quad (C \doteq 0.9)$$

なる評価式がえられている。(Berry, Esséen らによって)

一方, U統計量についても, 適当な仮定のもとで, それは漸近的に正規分布にしたがうことはよく知られているが, 上のような評価式はまだえられていないようである。

ここではチェビシェフ型の不等式と上の評価式を利用して, 同様な評価式を求めてみたい。ただし, 今のところ十分よい結果はえられておらず, 改良の余地があると思われる。

§ 2. 収束速度

ある分布関数族 $\mathcal{F} = \{F\}$ (以下に述べる仮定をみたす) について, 推定可能なパラメータ $\theta(F)$ とその不偏推定量 $g(X_1, \dots, X_m)$ を考える。いま大きさ $n (> m)$ の標本 $X_1, X_2, \dots, X_m$ からつくられる U 統計量

$$(2.1) \quad U_n = \binom{n}{m}^{-1} \sum_{(i_1, \dots, i_m)} g^*(X_{i_1}, X_{i_2}, \dots, X_{i_m})$$

を考える。ただし, 右辺の和は可能な組合せ $(i_1, i_2, \dots, i_m)$ にわたり, g^* は g を対称化したものとする ($g \in L_2(F)$)。

よく知られているように, U_n は θ の不偏推定量で, その分散は

$$(2.2) \quad V(U_n) = \frac{m^2}{n} \tau_1^2 + O(n^{-2})$$

ただし $\tau_1^2 = V[g_1^*(X_1)]$ かつ

$$g_1^*(x_1) = E\{g^*(X_1, X_2, \dots, X_m) \mid X_1 = x_1\}$$

であるから, U_n を規準化した Z_n :

$$(2.3) \quad Z_n = \frac{\sqrt{n}}{m\tau_1} (U_n - \theta)$$

を考えると, $E\{Z_n\} = 0$ かつ

$$(2.4) \quad V(Z_n) = 1 + \frac{C_1}{n} + O(n^{-2})$$

である。一方

$$(2.5) \quad Y_n = \frac{1}{\sqrt{n} \tau_1} \sum_{i=1}^n [g_1^*(X_i) - \theta]$$

を考えると, $E\{Y_n\} = 0$, $V(Y_n) = 1$ で, かつ $g_1^*(X_i)$ の3次の絶対積率 β_3' が存在すれば, (1.1) によって

$$(2.6) \quad \sup_{-\infty < y < \infty} |G_n(y) - \Phi(y)| < \frac{c\beta_3'}{\sqrt{n} \tau_1^3}$$

となる。また, $\text{Cov}(Y_n, Z_n) = 1$ で,

$$(2.7) \quad E\{(Y_n - Z_n)^2\} = \frac{C_1}{n} + O(n^{-2})$$

となることも容易に確かめられる。(2.7) より Y_n の分布 G_n と Z_n の分布 H_n との差を評価することができ (チェビシェフ型の不等式をもちい),

$$(2.8) \quad \sup_{-\infty < z < \infty} |H_n(z) - G_n(z)| \leq C_2 n^{-\frac{1}{3}} + O(n^{-\frac{4}{3}})$$

がえられる。(2.6) と (2.8) より

$$(2.9) \quad \sup_{-\infty < z < \infty} |H_n(z) - \Phi(z)| \leq C_2 n^{-\frac{1}{3}} + C_3 \beta_3' n^{-\frac{1}{2}} + O(n^{-\frac{4}{3}})$$

なる評価式がえられる。

参考文献

- [1] Berry, A. C. (1941): The Accuracy of the Gaussian Approximation of the Sum of Independent Variables, Trans. Amer. Math. Soc., Vol. 49, pp. 122-136.
- [2] Esséen, C. G. (1945): Fourier Analysis of Distribution Functions; A Mathematical Study of the Laplace-Gaussian Law, Acta Math., Vol. 77, pp. 1-125.
- [3] Esséen, C. G. (1956): A Moment Inequality with an Application to the Central Limit Theorem, Skand. Aktuar. Tidskr., pp. 160-170.
- [4] Fraser, D. A. S. (1957): Nonparametric Methods in Statistics, John Wiley, pp. 223-228.