関数挙動に対する離散化の影響

志村　隆彰　数理・推論研究系 准教授

1　研究の背景・動機

連続分布を離散化したときの影響に関心がある。具体的には、確率論の様々な極限定理と関連が深い、確率分布の裾 或いは 切断積率の漸近挙動 ($x \to \infty$) が離散化によりどのような影響を受けるか、あるいは受けないか、実質的には、べき関数を一般化した正則変動関数と関連する関数に対する離散化の影響を考察する。

$$1 - F(x), \quad \int_{|t|<x} t^{\rho} F(dt) \ (\rho \geq 0).$$

例１：　分布 F に従う i.i.d. 確率変数列 $\{X_k\}$ の部分最大値 $Z_n = \max\{X_k : 1 \leq k \leq n\}$ は、指数分布のように $1 - F(x) \sim e^{-x}$ であれば、適当な定数列 $a_n > 0$ と $b_n \in \mathbf{R}$ によって、$(Z_n - b_n)/a_n$ の分布をグンベル分布 $\exp(-e^{-x})$ に収束させることができる。一方、F が指数分布の離散分布版といえる幾何分布の場合、離散化によって裾$1 - F(x)$ の漸近挙動が変ってしまい、収束させられない。

例２：　正値分布 Fの切断平均 $\int_0^x tF(dt)$ が緩慢変動（後述）すれば、F に従う i.i.d. 確率変数列の部分和　$S_n = \sum_{k=1}^n X_k$ が相対安定、すなわち、ある定数列 $\{c_n : n \in \mathbf{N}\}$ により、

$$S_n/c_n \to 1 (n \to \infty) \quad \textit{in probability}$$

とできる（逆も正しい）。たとえば、片側コーシー分布$p(t) = 2\frac{c}{\pi}\frac{1}{(t-\gamma)^2+c^2}$ では、

$$\int_0^x tp(t)dt \sim 2\frac{c}{\pi}\log x \qquad x \to \infty$$

となる。しかし、裾挙動はコーシー分布と異なる、以下の離散分布ペーター・ポール分布の切断平均も切断平均は対数オーダーとなる。

$$P(X = 2^k) = 2^{-k} \quad (k = 1, 2, \cdots),$$

このように切断平均が対数オーダーとなる分布の裾には多様性がある。

注意: 離散化は、実データの粗さを表しているとみなせるから、統計的には、データの粗さの影響の考察といえる。

2　問題：数列による関数の離散化

以下では、問題のキーとなる関数の離散化について述べる。

【数列による関数の離散化】$[0, \infty)$ 上の単調連続関数 $f(x)$ と単調で発散する増加正値数列 $\{a_n > 0 : n \in \mathbf{N}\}$に対して、関数 f の 数列 $\{a_n\}$ による離散化（関数）$g(x)$ を

$$g(x) = f(a_n) \quad for \ \ x \in [a_n, a_{n+1})$$

で定義し、主に次のようになる場合を考える。

$$\lim_{x\to\infty} f(x)/g(x) = 1 \ (f(x) \sim g(x)).$$

f の離散化を f_d で、特に、$a_n = n$ のとき、f_1 で表す。次が成り立つことに注意する。

【命題０】任意の単調関数 f に対して、$\lim_{n\to\infty} a_n = \infty$ となる正数列 $\{a_n : n \in \mathbf{N}\}$ で、$f_d \sim f$ となるものが存在する。

次の問題を考える。

【問題】

1．$f_1 \sim f$ となる f を特徴づける。
2．与えられた f に対して、$f \sim f_d$ となる $\{a_n\}$ の条件を決める。
3．階段関数 g に対して、$f_d = f$ となる”良い”連続関数 f を求める。
4．f がもつ性質が f_d に遺伝するか。

3　正則変動関数及び関連する関数

考察の対象となる関数を定義する。

【正則変動関数】ある半開区間$[x_0, \infty)$上の正値関数$f(x)$が指数 $\rho(\in \mathbf{R})$ の正則変動関数(regularly varying function)であるとは、任意の$k > 0$に対して、

$$\lim_{x\to\infty} \frac{f(kx)}{f(x)} = k^{\rho}$$

が成り立つときをいう。$f \in \mathbf{RV}_{\rho}$ とかく。特に、$\rho = 0$ のとき、緩慢変動関数(slowly varying function)といい、指数ρの正則変動関数は、緩慢変動関数 $l(x)(\in \mathbf{RV_0})$ により、

$$f(x) = x^{\rho} l(x)$$

と書ける。

例３：　$f(x) = c(c > 0)$、$\log x$、$(\log x)^2$、$\exp(\sqrt{\log x})$.

【Γ変動関数】$[x_l, x_0)(x_0 \leq \infty)$ 上の正値非減少関数 f がΓ変動するとは、$\lim_{x\to x_0} f(x) = \infty$ かつ、$[x_l, x_0)$ 上の正値関数 $s(x)$ により、任意の $k \in \mathbf{R}$ に対して、

$$\lim_{x\to x_0} \frac{f(x + ks(x))}{f(x)} = e^k$$

が成り立つときをいう。$f \in \mathbf{\Gamma}$ とかく。

例４：$f(x) = \exp x$ は $s(x) = 1$としてΓ変動関数である。

【Π変動関数】$[x_l, \infty)$ 上の正値非減少関数 f がΠ変動するとは、任意の $k \in \mathbf{R}$ に対して、

$$\lim_{x\to\infty} \frac{f(kx) - f(x)}{f(ex) - f(x)} = \log k$$

が成り立つときをいう。$f \in \mathbf{\Pi}$ とかく。

例５:$f(x) = \log x$ は Π変動関数である。

注意: Γ変動関数とΠ変動関数は互いに逆関数の関係にある。

4　離散化された関数の性質

【命題１】$f \in \mathbf{RV}_{\rho}(\rho \neq 0)$、$f$ の数列 $\{a_n\}$ による離散化 $f_d(x)$ とするとき、$f \sim f_d$ となるための必要十分条件は、

$$\lim_{n\to\infty} a_{n+1}/a_n = 1$$

である。 特に、$f_1 \sim f$ となる。

【命題２】$f \in \mathbf{RV}_0$ ならば、

$$\limsup_{n\to\infty} a_{n+1}/a_n < \infty$$

を満たす任意の数列 $\{a_n\}$ に対して、

$$\lim_{n\to\infty} f(a_{n+1})/f(a_n) = 1.$$

すなわち、$f \sim f_d$ が成り立つ。

【命題３】すべての $f \in \Gamma$ に対し、$f_1 \notin \Gamma$ である。

例６：　$e^x \in \Gamma$ だが、$\lfloor e^x \rfloor \notin \Gamma$.

【命題４】$f \in \mathbf{\Pi}$ に対し、$f_1 \in \mathbf{\Pi}$ となるための必要十分条件は、$\lim_{x\to\infty} f(ex) - f(x) = \infty$ である。

例７：$\log x, (\log x)^2 \in \Pi$ に対し、$\lfloor \log x \rfloor \notin \Pi$, $\lfloor (\log x)^2 \rfloor \in \Pi$ である。