論文の内容の要旨

論 文 題 目：Asymptotic Inference for Stochastic Differential Equations with Jumps from Discrete Observations and Some Practical Approaches

（飛躍型確率微分方程式に対する離散的観測に基づく漸近推測理論，及びその実際的方法）

氏　　名：　清 水　泰 隆

本論文では，以下のような飛躍型確率微分方程式に従う連続時間型確率過程 $X := \{X_t\}_{t\geq 0}$ に対する統計的漸近推測論を扱う．

$$X_t = X_0 + \int_0^t a(X_s)\,ds + \int_0^t b(X_s)\,dw_s + \int_0^t \int_{\mathcal{E}} c(X_{s-}, z)\,(p-q)(ds, dz) \tag{1}$$

ここに，$\mathcal{E} = \mathbb{R}^d \setminus \{0\}$，$w$ は多次元ウィナー過程，p は $[0,\infty)\times\mathcal{E}$ 上のポアソンランダム測度，q はその補正測度であり，$q(ds, dz) = f(z)\,dzds$ なる表現を持つとする．f は時間依存しないレヴィ密度である．係数 a, b, c はそれぞれ適切な次元を以って定義された可測関数である．このような確率過程 X は，飛躍型拡散過程と言われる．

飛躍型拡散過程 X をある現象の確率モデルとして用いる場合，一般には，関数 a, b, c やランダム測度 p の特徴付けとなるレヴィ密度 f などは未知であり，X の観測からそれらを統計的に推定することは，応用上重要な問題である．

ここで，X の観測には大別して 2 種類考えられる．一つは連続的観測であり，これは X のパスを観測期間を通して完全に観測したものである．もう一つは離散的観測であり，これは観測期間におけるいくつかの離散時点のみの X の値を観測したものである．連続的観測の下での統計推測は，数理統計的には重要な問題の一つではあるが，実際のデータを扱う場合の方法論としてはあまり現実的とはいえない．そこで，我々は連続時間型の確率過程 X の離散的観測による統計推測に興味をもつことになる．本論文の主題は，このような離散的観測に基づく X の統計的推測である．

論文を通して，X の離散的観測は初期値 X_0 のほかに n 個の標本として与えられるとし，観測の時間幅は標本数 n に依存して等間隔 $h_n\,(>0)$ とする．すなわち，離散的観測は時刻 $t_i^n = i\times h_n\ (i=0,1,2,\dots,n)$ においてなされると仮定する．このとき我々の目的は，標本 $X^n := \{X_{t_i^n}\}_{i=0,1,\dots,n}$ から a,b,c，及び f に含まれ得る未知量に対する推定量を構成し，n が増大するときの推定量の漸近挙動を調べることである．

このような離散的観測に基づく漸近理論を論ずるとき，観測幅 h_n の n に対する漸近挙動（観測スキーム）を定めておくことは，推定量の漸近分布に影響を与えるという点で本質的に重要である．本論文では $n\to\infty$ のとき，$nh_n\to\infty$，更に，ある定数 $\delta>0$ に対して $nh_n^{1+\delta}\to 0$ となることを仮定して議論する．定数 δ は観測の頻度を制御し，δ が大きいほど観測の頻度が低いことを許容する．この値が推定量の漸近効率に影響を与えるという点で，δ の上限は離散観測に基づく漸近理論の興味の対象の一つであるが，本論文においては，推定量の漸近効率を優先して $\delta\in(0,1]$ の下で議論しており，実際にいくつかの場面で推定量の漸近有効性が達成される．

本論文は主に６つの章から成る．このうち，第１章では統計的漸近理論の歴史的経緯と本論文の概要を簡単に述べることで序章とし，第２章で飛躍型拡散過程の定義といくつかの性質について述べている．主題となる統計的推測論は，第３章から第６章に渡って議論される．

第３章では，関数 a,b,c，及びレヴィ密度 f が，それぞれ多次元の未知パラメータ θ,σ を含む既知の関数 $a(x,\theta),b(x,\sigma),c(x,z,\theta)$，及び $f_\theta(z)$ として与えられている場合に，パラメータ (θ,σ) の同時推定（パラメトリック推定）を論じている．以下，$\alpha:=(\theta,\sigma)$ とし，それらの真値を α_0,θ_0，及び σ_0 と表すことにする．

この章では，レヴィ密度 f_θ に対して，

$$\lambda_0 := \int_{\mathcal{E}} f_{\theta_0}(z)\,dz < \infty \tag{2}$$

となることを仮定する．このようなモデルを仮に「ポアソン型」と呼ぶことにする．また，推定量の漸近挙動を調べる上で本質的な極限定理を得るために，X のエルゴード性を仮定する．

パラメトリック推定における有力な推定量の構成法の一つに最尤推定法があるが，離散観測 X^n の推移確率を陽に表現することは困難であるため，X^n の尤度関数を直接書き下すことは難しい．そこで本論文では，X の飛躍を標本 X^n から検出し，その飛躍の大きさを近似することによって X^n の対数尤度関数の近似を求めており，この飛躍の検出法こそが本論文全体を通しての鍵となる．

仮定 (2) の下では，n が大きい時，したがって h_n が十分小さい時，各区間の上で起こる飛躍の回数は高々1 回と近似することが出来る．もし，ある区間 $(t_{i-1}^n, t_i^n]$ の上で飛躍が起こらなければ，X はその間を拡散過程 $dX_t = a(X_t,\theta_0)\,dt + b(X_t,\sigma_0)\,dw_t$ に従って推移するので，直感的には，時間 h_n における増分 $|\Delta_i X^n| := |X_{t_i^n} - X_{t_{i-1}^n}|$ は小さく，$h_n\to 0$ のときその増分は 0 に近づく．しかし，１回の飛躍があった場合，その増分は確実に飛躍幅に近づくであろう．そこで，飛躍の基準となる閾値 r_n を n に応じて適切に定めることによって，$|\Delta_i X^n| > r_n$

となる区間では飛躍があったと推測し，その大きさを $\Delta_i X^n$ によって近似するのである．

このようにして飛躍の無い区間とある区間を推測し分類しておく．飛躍が無いと推測される区間上では，拡散過程の推移近似である「局所ガウス近似」を用いて尤度を近似し，飛躍があると推測される区間上では，ポアソンランダム測度に対する尤度の自然な離散化を用いることによって X^n の近似尤度を構成することができ，これらを α_0 の推定関数として採用するのである．実際，この推定関数による推定量 $\hat{\alpha}_n = (\hat{\theta}_n, \hat{\sigma}_n)$ に対して，$\sqrt{nh_n}(\hat{\theta}_n - \theta_0)$ と $\sqrt{n}(\hat{\sigma}_n - \sigma_0)$ はそれぞれ漸近正規し，ある意味で漸近有効な推定量となる．ここで，$\hat{\theta}_n$ と $\hat{\sigma}_n$ で最適な収束率が異なっていることに注意が必要で，これが α を θ と σ に分けて表現した所以である．

既述したように，漸近有効性を得るには観測スキームにおける定数 δ の定め方が重要であるが，第３章ではレヴィ密度 f にある種の正則条件を与え，$\delta = 1$ の下での漸近有効性を達成しており，このスキームは，この種の漸近論における標準的な設定である．f に対する条件を緩和させることは可能であるが，その場合には $\delta < 1$ とする必要が生じ，どちらを選ぶかは応用の対象と観測の状況によって異なるであろう．

第４章においては，真のレヴィ密度に対して，

$$\int_{\mathcal{E}} f_{\theta_0}(z)\, dz = \infty \tag{3}$$

となる場合も一部許容した上でパラメトリック推定を行っている．このモデルを「レヴィ型」と呼ぶことにする．X のエルゴード性は引き続き仮定する．

条件 (3) の状況では，観測幅 h_n の区間上で常に無限個の飛躍が存在することになり，上述のような判定は難しいように見える．しかし，確率微分方程式 (1) の飛躍項（右辺第４項）を，ある数列 $\varepsilon_n > 0$ に対して，以下の２項

$$\begin{aligned} B_n(t) &:= \int_0^t \int_{0<|z|<\varepsilon_n} c(X_{s-}, z, \theta_0)\,(p - q_{\theta_0})(ds, dz) \\ J_n(t) &:= \int_0^t \int_{|z|>\varepsilon_n} c(X_{s-}, z, \theta_0)\,(p - q_{\theta_0})(ds, dz) \end{aligned}$$

の和に分解すれば，上記 $J_n(t)$ から生じる飛躍の個数は観測幅 h_n の区間上で常に有限個であり，また，ε_n が減少する時，$B_n(t)$ は適当な正則条件の下で拡散項とある意味類似の挙動をすることが示される．これはある意味で，レヴィ型モデルのポアソン型モデルによる近似とみることも出来るであろう．このような観点に立つと，既述した飛躍判別法が一部適用できて，第３章と類似の議論が可能になる．ただし，飛躍ごとの検出は困難で，飛躍に関わるパラメータの漸近有効性は放棄せざるを得ない．また，ε_n，h_n，及び r_n らの関係，あるいは，それらと観測スキームにおける定数 δ との間の関係を注意深く定める必要がある．

第５章では，再びポアソン型 (2) に戻り，レヴィ密度 f のノンパラメトリック推定について考察している．ここでは a, b や f を未知とし，パラメトリックな構造を仮定しないが，c は既知の関数と仮定して議論する．

今，連続的観測 X が得られていたと仮定すると，X の全ての飛躍時刻とその大きさを観測できることになる．このとき，任意の x に対して，関数 $z \mapsto y = c(x,z)$ に逆関数 $y \mapsto z = c^{-1}(x,y)$ が存在したとして，ある飛躍時刻を τ とすると，$\Delta z_\tau := c^{-1}(X_{\tau-}, \Delta X_\tau)$ は既知である．ただし，$\Delta X_\tau := X_\tau - X_{\tau-}$ とする．このとき，Δz_τ は確率密度 $\lambda_0^{-1} f(z)$ を持つ確率変数と見なせるので，統計学において古典的な密度推定法を応用することにより f のカーネル型推定量を構成できる．離散的観測の場合には，第 3 章で用いた飛躍判別の議論によって，$|\Delta_i X^n| > r_n$ となる区間における飛躍 Δz_τ の自然な近似として $\Delta_i z^n := c^{-1}(X_{t_{i-1}^n}, \Delta_i X^n)$ を用いることにより，連続的観測の場合の推定量の自然な離散化としてカーネル型推定量を構成することができる．本章では，そのカーネル型推定量の平均 2 乗誤差の意味での一致性を示し，その最適な収束率と誤差限界を求めている．

第 5 章の方法は，第 3 章で本質的であったエルゴード性の仮定を排除したという点において応用上有力である．また，第 3 章で仮定した f の正則条件も若干緩和して議論されている．ただし，既述したように，f の条件の緩和によって観測スキームは $\delta \in (0, 1/2)$ と制限的にならざるを得ない．

さて，これらの議論を実際に応用する場合，飛躍判別の閾値 r_n の決め方が最も重要であることは言うまでもない．しかし，漸近理論が与えてくれる閾値 r_n の決定条件は，標本数 n が増大する時の漸近挙動だけであり，実際，閾値に関する一定の収束条件を満たす数列 r_n を取れば，その定数倍 Lr_n $(L > 0)$ もまた飛躍判別の閾値として漸近的には同等な役割を果たす．しかし，実際には標本数 n は固定されているし，論文内で示したいくつかの数値実験が示すように，L の値によって推定精度は劇的に変化する．そこで，n や真のモデルに応じた r_n の選択が重要になる．これに対して，第 5 章でいくつか直感的に有効と思われる r_n の選択法を示すが，第 6 章において，離散的観測による r_n の機械的な選択のアルゴリズムを提案し，その数理的正当性について論じている．

ここで提案される手法は，第 5 章の直感的手法よりも簡便であり，一組の観測が与えられればそれに対する r_n は一意的に決定されるという意味で，実用性の高い手法と言えるであろう．ただし，ここでのモデル設定は $c(x,z) \equiv z$ に限っており，いまのところ限定的と言わざるを得ない．更に一般的なモデルに対する選択法の開発は，将来に対する重要な課題であろう．