確率過程

確率実験の結果を表すものが確率変数であるが、コイン投げを続けて投げる独立試行とか、時刻とともに変化する事象を観測するばあいには、結果が繰り返しの回数や時刻に依存して変化する。一般に時間パラメータ t に伴って起こる変化の結果（確率変数）を $X(t)$ あるいは X_t と表し、確率過程とよぶ。コイン投げの回数とその結果の集計値、株式市場の株価変動、ある個体の生存数変化、製造機械による製品の規格状態など、身近にはたくさんのものが考えられる。時間パラメータには整数による離散型確率過程と実数による連続型確率過程がある。また各時間 t で変数のとる値 $X(t)$ の集合 (非負の整数や実数の集合) を状態空間といい、任意数 n での区分点 $t_0 < t_1 < t_2 < \cdots < t_n$ において増分を表す確率変数 $X(t_1) - X(t_0), X(t_2) - X(t_1), \cdots X(t_n) - X(t_{n-1}),$ が独立であるとき、独立増分をもつという。

代表的な確率過程の例

２項過程 （時間間隔は幾何分布、その和は負の２項分布）独立増分で定常な推移確率をもつ離散形時間の確率過程 $\{X_i; i = 1, 2, \cdots, \}$ で、パラメータ p のベルヌーイ試行（コイン投げ）の列を考える。いま $X_1 = 0$ とし、最初に表が出るまでに必要とした時間を X_2、つぎに再び表が出るまでの時間 X_3 と順次定める。つまり X_n は時刻 n までに成功した（コインが表が出た）事象の回数を表す。これを２項過程という。表が出ることの間隔時間は幾何分布にしたがい、これを $t_1, t_2, \cdots$ とすれば、$P(X(t_{i+1}) = k | X(t_i) = j) = \begin{cases} 1,\, k = j+1 \\ 0,\, k \neq j \end{cases}$ で、一般には２項分布の値と同じ確率となります。

$$P(X(n) = k | X(1) = 0) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \cdots, n$$

ポアソン過程 （時間間隔は指数分布、その和がガンマ分布）上述の離散時刻を連続時間において考えたタイプ、すなわちある事象の起こる回数を数え上げること、計数過程がポアソン過程である。あるコンビ二・ストアーに到着する客の人数、ある交差点での交通事故の発生件数、ある多数部品から構成されるシステムにおける故障する部品の個数などは計数過程の例と考えられる。起こりにくい事象を多数回繰り返すことの結果に注目する。起こる起こらないことはコインの表が出る出ないことであり、この回数は２項分布で計算できる。繰り返し数を多くし、生起する確率が非常に小さいならばポアソン分布に近づく。同様に時間間隔を表す幾何分布は指数分布へ、またある時刻までの総生起回数を表す負の２項分布はガンマ分布へと近づく。これらから、２項過程の連続時間版がポアソン過程とよばれる由縁である。

$\{X_t; t \geq 0\}$ はパラメータ λ のポアソン過程であるとは独立増分な過程で

$$P(X_h = j \,|\, X_0 = 0) = \begin{cases} 1 - \lambda h + o(h), & j = 0 \\ \lambda h + o(h), & j = 1 \\ o(h) & j \geq 2 \end{cases}$$

マルコフ連鎖 $\{X_n; n = 0, 1, 2, \cdots\}$ における状態空間が非負整数値 $\{0, 1, 2, \cdots\}$ をとり、つぎの推移確率とよばれる条件付き確率の性質をもつとき、マルコフ連鎖という。すなわち次ステップ X_{n+1} の運動法則が現時点での条件 X_n のみに依存して過去の $X_{n-1} = i_{n-1}, \ldots, X_0 = i_0$ にはよらない場合をさす：$P\{X_{n+1} = j \,|\, X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0\} = P_{ij}$ このとき (1) $P_{ij} \geq 0, \forall i, j;$ (2) $\sum_{j=0}^{\infty} P_{ij} = 1, \forall i$ の性質を満たす。推移確率（あるいは推移行列）

とは

$$\boldsymbol{P} = (P_{ij}) = \begin{pmatrix} p_{00} & p_{01} & p_{02} & \cdots \\ p_{10} & p_{11} & p_{12} & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix}$$

をいう。行列の積に対応した演算を満たす ; $P^{n+m} = P^m P^n = P^n P^m$, つまり $p_{ij}^{n+m} = \sum_k p_{ik}^m p_{kj}^n = \sum_k p_{ik}^n p_{kj}^m$ である。また $P(X_0 = i), i \in \{0, \pm 1, \pm 2, \cdots\}$ を**初期確率**と定義する。

定義 ;（状態の分類）状態 i が **再帰的** (recurrent) とは $\sum_{n=1}^{\infty} p_{ii}^n = \infty$，状態 i が **過渡的** (transient) とは $\sum_{n=1}^{\infty} p_{ii}^n < \infty$。

具体的な例としては

「ランダムウォーク（酔歩)」ある適当な実数 $0 < p < 1$ について状態空間 $\{i = 0, \pm 1, \pm 2, \}$ における推移確率 $P_{i,i+1} = p$, $P_{i,i-1} = 1 - p$, $P_{i,j} = 0$ otherwise で与えられる。

「ギャンブラー破産の問題」状態空間を $\{1, 2, \cdots, N\}$ として、これをギャンブルでの所持金額を表すとします。勝ち負けで 1 単位ずつ金額が変動して前述の「ランダムウォーク」と同じに変化 $P_{i,i+1} = p$, $P_{i,i-1} = 1 - p$, $P_{i,j} = 0$ すると考えます。しかし破産が $\{0\}$ で起こり、これ以上は賭けを続けることができませんから、$P_{00} = P_{NN} = 1$ という吸収壁をもちます。

マルコフ過程 連鎖と過程の違いは主にとり得る値が離散か連続かによることが多い。つぎに説明することは、連続時間のマルコフ連鎖と考えたほうがより正確な用語の定義であるが、マルコフ過程はこれをより一般化したと考える拡散過程をイメージしてもらいたい。連続時間において状態空間が整数の集合である、出生死滅過程を典型例として述べる。ポアソン過程では、事象の起こった回数に注目するから、順次に 1 つずつ増加する。これに対して、増加と減少も同時にありえるものとして考えた場合である。確率過程 $\{X(t); t \geq 0\}$ において推移確率は $p_{ij}(t) = P(X(s+t) = j \mid X(s) = i), \quad t \geq 0, i, j \in S$ ここで S は状態空間を表す集合とする。またマルコフと呼ばれる由縁は将来の運動法則が現在の状態のみにより定まり、過去の状態には依存しない、ということである。これは指数分布のメモリーレスに直接関係している。

出生死滅過程とは状態空間 $S = \{\cdots, -1, 0, 1, 2, 3, \cdots\}$ で

$$P(X(h) = j \mid X(0) = i) = \begin{cases} \lambda_i h + o(h) & j = i + 1, \\ 1 - (\lambda_i + \mu_i)h + o(h) & j = i, \\ \mu_i h & j = i - 1 \\ o(h) & otherwise \end{cases}$$

ブラウン運動（拡散過程） 気体や液体などの溶媒中に溶けた物質は、濃度の高い方から低い方へと拡散していき、その単位時間当たりの拡散量は、その時点での濃度勾配に比例する。これと同じく熱の伝導も同じ原理で行われる。1 次元では $u = u(t, x)$ を時刻 t における点 x での温度を表すとする。$\dfrac{\partial u}{\partial t} = \dfrac{1}{2} \dfrac{\partial^2 u}{\partial x^2}$ これが熱方程式とよばれる。$\lim_{t \to 0} u(t, x) = f(x)$ とおくと、この解は $u(t, x) = \int_{R^1} g(t, y) f(x - y) dy$ ただし $g(t, x) = \dfrac{1}{\sqrt{2\pi t}} e^{-\dfrac{x^2}{2t}}$ この $g = g(t, x)$ も上の熱方程式を満たすことがわかる。熱核とよばれる。これは正規分布 $N(0, t)$ の密度関数に他ならない。

$g(t, x) = \dfrac{d}{dx} P(X(t) \leq x)$

連続時間パラメータの拡散過程 $\{X(t); t \geq 0\}$ とは (1) 独立増分 (2) 連続で過程であり、$X(0) = 0$ (3) 増分が平均ゼロ、分散が時間間隔をもつ正規分布にしたがう、つまり $X(t+s) - X(s)$

は $X(s)=y$ には依存せず、$N(0,t)$ にしたがう（時間間隔 t を分散とすることに注意）。
$E[X(t+s)-X(s)]=0, \quad E[(X_{t+s}-X_s)^2]=t$

演習問題

$\boxed{1}$ ２つの状態 $\{0,1\}$ からなる時間一様なマルコフ連鎖の推移行列が $P:=\begin{pmatrix}0.7 & 0.3\\ 0.4 & 0.6\end{pmatrix}$ であり，初期確率が $P(X_0=0)=0.4, P(X_0=1)=0.6$ で与えられている。このとき，つぎの確率を計算せよ。

(1) $P(X_2=1|X_0=1)$ (2) $P(X_3=0|X_1=0)$
(3) $P(X_4=0)$ (4) $P(X_4=1,X_2=1|X_3=1)$

$\boxed{2}$ ４つの状態空間 $\{0,1,2,3\}$ における時間一様なマルコフ連鎖の推移確率が $P:=\begin{pmatrix}0 & 0 & 1/2 & 1/2\\ 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 1 & 0 & 0\end{pmatrix}$，であるとき，どの状態が，過渡的かあるいは再帰的であるか？

$\boxed{3}$ ギャンブラー破産の問題で初期所持金 i からはじめていつかは破産する確率 $P_i=P(\exists N; X_N=0|X_0=i)$ は

(1) $\quad P_i=pP_{i+1}+qP_{i-1}$ (2) $\quad P_i=\begin{cases}\dfrac{1-(q/p)^i}{1-(q/p)^N} & \text{if} \quad p\neq 1/2\\ \dfrac{i}{N} & \text{if} \quad p=1/2\end{cases}$

(3) $\quad N\to\infty$ とすると $P_i\to\begin{cases}1-(q/p)^i, & \text{if} \quad p\neq 1/2\\ 0, & \text{if} \quad p=1/2\end{cases}$

となることを示せ。ヒント：$P_{i+1}-P_i, P_i-P_{i-1}$ の関係式は？

$\boxed{4}$ 酔歩 $p=1/2$ のとき、$P(X_n=x|X_0=0)=\dfrac{1}{2^n}\begin{pmatrix}n\\ \frac{n+x}{2}\end{pmatrix}, x=0,1,2,\cdots$ （ただし $\dfrac{n+x}{2}$ は $0,1,\cdots,n$ とする）を示せ。ヒント：２項定理をもちいればよい。

$\boxed{5}$ ２項過程からポアソン過程への導きを、幾何分布と指数分布、負の２項分布とガンマ分布の対応から示しなさい。

$\boxed{6}$ X_t をブラウン運動とするとき、$Y(t)=cX(t/c^2)$ も独立増分をもち、連続かつ、増分が正規分布にしたがう。この正規分布の平均と分散を計算しなさい。これも広い意味でのブラウン運動とよばれる。