定常ポアソン過程

@phykm

2018 年 11 月 25 日

概要

ポアソン過程の構成と特徴づけについて。とある本を読んでいたときに、停止時間差分の独立性と指数分布であることの証明が、「直感的説明」によってスルーされており、大変悩んで拘泥したために、これを記す。

1 指数分布

Definition 1.1. 実数ボレル族上の確率測度で、

$$P([0,t]) = 1 - \exp(-\lambda t) \tag{1}$$

であるものを、パラメータ λ の指数分布という。ある実数値確率変数が指数分布であるとは、それの押し出しが指数分布であることをいう。あきらかにこのような変数は $X \geq 0, a.s.$ である。

Proposition 1.2. パラメータ λ 指数分布に従う確率変数 X について、期待値は λ^{-1}、分散は λ^{-2}

Proof. (ルベーグ測度に対する) 確率密度関数は $\lambda \exp(-\lambda t)$ であるから、特性関数が

$$\int_0^\infty \exp(-\lambda t + ixt)dt = \frac{1}{1 - ix\lambda^{-1}} \tag{2}$$

これの x 展開 1 次は λ^{-1}、2 次は λ^{-2}、よって、期待値は λ^{-1}、二次モーメントは $2\lambda^{-2}$、分散は $2\lambda^{-2} - \lambda^{-2} = \lambda^{-2}$

Proposition 1.3. X がパラメータ λ の指数分布に従うとする。このとき

$$P(X > t + s | X > t) = P(T > s) \tag{3}$$

なお、不等号を等号含むものにしても同様である。

Proof.

$$\begin{aligned} & P(X > t + s | X > t) && (4) \\ =& \frac{P(X > t + x \wedge X > t)}{P(X > t)} && (5) \\ =& \frac{\exp(-\lambda(t + s))}{\exp(-\lambda t)} && (6) \\ =& \exp(-\lambda s) && (7) \\ =& P(X > s) && (8) \end{aligned}$$

Proposition 1.4. X_i がパラメータ λ_i の指数分布従う独立な有限個の確率変数であるとする。このとき、$\min_i X_i$ は $\sum_i \lambda_i$ の指数分布に従う。

Proof.

$$P(\min_i X_i > t) = \Pi_i P(X_i > t) = \exp(-\sum_i \lambda t) \tag{9}$$

Proposition 1.5. X_i がパラメータ λ_i の指数分布従う独立な n 個の確率変数であるとする。このとき、同じ空間に $\{1 \dots n\}$ に値をとる有限離散確率変数 I を、$I = \mathrm{argmin}_i X_i$ で作る。この確率分布は

$$P(I = i) = \frac{\lambda_i}{\sum_j \lambda_j} \tag{10}$$

である。

Proof. $n = 2$ の時に示し、前命題を使って帰納する。$n = 2$ のとき、

$$P(I = 2) = P(X_1 > X_2) \tag{11}$$
$$= \int_{x>y>0} \lambda_1 \lambda_2 \exp(-\lambda_1 x - \lambda_2 y) dxdy \tag{12}$$
$$= \int_0^\infty (1 - \exp(-\lambda_2 x))\lambda_1 \exp(-\lambda_1 x) dx \tag{13}$$
$$=1 - \frac{\lambda_1}{\lambda_1 + \lambda_2} \tag{14}$$
$$=\frac{\lambda_2}{\lambda_1 + \lambda_2} \tag{15}$$

$n = k$ のときに題意とする。

$$P(I = n+1) = P(\min_{i=1\dots n} X_i > X_n + 1) \tag{16}$$
$$=\frac{\lambda_{n+1}}{\sum_{i=1\dots n+1} \lambda_i} \tag{17}$$

であるから帰納される。

Proposition 1.6. 以上の 2 つ $V = \min_i X_i$ と I は独立である。

Proof.

$$P(I = i, V > t) \tag{18}$$
$$=P(I = i, X_j > t(\forall j)) \tag{19}$$
$$=P(X_i < X_j(\forall j \neq i)) \tag{20}$$
$$= \int_0^t \lambda_i \exp(-\lambda_i t) \Pi_{j \neq i} \exp(-\lambda_j t) dt \tag{21}$$
$$=\frac{\lambda_i}{\sum_j \lambda_j}(1 - \exp(-\sum_j \lambda_j t)) \tag{22}$$
$$=P(I = i)P(V > t) \tag{23}$$

Proposition 1.7. $\{X_i\}_{i=1}n$ をいずれも独立なパラメータ λ の指数分布とする。$\sum_i X_i$ は次のガンマ分布に従う。この確率密度関数は

$$p(t) = \lambda \exp(-\lambda t)\frac{(\lambda t)^{n-1}}{(n-1)!} \tag{24}$$

Proof. 帰納する。$n=1$ は明らかに正しい。$n=k$ で成り立つとする。

$$\int_{x+y<s} \lambda \exp(-\lambda x)\lambda \exp(-\lambda y)\frac{(\lambda y)^{n-1}}{(n-1)!}dxdy \tag{25}$$

$$= \int_{0<y<s} (1-\exp(-\lambda(s-y)))\lambda \exp(-\lambda y)\frac{(\lambda y)^{n-1}}{(n-1)!}dy \tag{26}$$

$$= \int_0^s \lambda \exp(-\lambda x)\frac{(\lambda x)^n}{(n)!}dx \tag{27}$$

2　指数分布からポアソン過程へ

Definition 2.1. $\{X_i\}_i$ は独立な可算個のパラメータ λ の指数分布とする。$T_0 = 0, T_n = \sum_{i=1}^n X_i$ としたときの、$N(t) = \max\{n : T_n \leq t\}$ を（パラメータ λ の）ポアソン過程と呼ぶ。

Proposition 2.2. ポアソン過程 $N(t)$ は、各時刻で $t\lambda$ のポアソン分布に従う。すなわち、

$$P(N(t)=n) = \exp(-t\lambda)\frac{(\lambda t)^n}{n!} \tag{28}$$

Proof. T_n はガンマ分布とわかっており、T_n と X_{n+1} は明らかに独立だから、X_{n+1} の確率について、T_n の値で次のように積分する。

$$P(N(t)=n) \tag{29}$$

$$=P(T_n \leq t \wedge X_{n+1} + T_n > t) \tag{30}$$

$$= \int_{x\geq t, y+x>t} \lambda \exp(-\lambda y)\lambda \exp(-\lambda x)\frac{(\lambda x)^{n-1}}{(n-1)!}dxdy \tag{31}$$

$$= \int_{x\geq t} \exp(-\lambda t)\frac{(\lambda x)^{n-1}}{(n-1)!}dx \tag{32}$$

$$= \exp(-t\lambda)\frac{(\lambda t)^n}{n!} \tag{33}$$

Proposition 2.3. ポアソン過程は独立増分であり、増分は $dt\lambda$ のポアソン分布に従う。

Proof. 　まず、$N(t+s)-N(s), N(s)$ の独立性を示す。$N(s)=n$ 条件付き $N(t+s)-N(s)$ の確率が、n

によらなければいい。指数分布 X_{n+1} の無記憶性を用いる。

$$
\begin{aligned}
&P(N(t+s)-N(s)<l|N(s)=n) && (34)\\
=&\frac{P(X_{n+1}+\dots X_{n+l}>s+t-(X_1+\dots X_n),X_{n+1}>s-(X_1+\dots X_n)\geq 0)}{P(X_n>s-(X_1+\dots X_n)\geq 0)} && (35)\\
=&\frac{\int_{x_1+\dots x_n\geq s}(\Pi_{i=1,\neq n+1}^{n+l}dP(x_i))P(X_{n+1}>s+t-(\sum_{i=1,\neq n+1}^{n+l}x_i),X_{n+1}>s-(\sum_{j=1}^{n}x_j)\geq 0)}{\int_{x_1+\dots x_n\geq s}(\Pi_{i=1}^{n}dP(x_i))P(X_{n+1}>s-(\sum_{j=1}^{n}x_j)\geq 0)} && (36)\\
=&\frac{\int_{x_1+\dots x_n\geq s,x_{n+2}+\dots x_{n+l}\leq t}(\Pi_{i=1,\neq n+1}^{n+l}dP(x_i))P(X_{n+1}>t-(\sum_{j=n+2}^{n+l}x_j))P(X_{n+1}>s-(\sum_{j=1}^{n}x_j)\geq 0)}{\int_{x_1+\dots x_n\geq s}(\Pi_{i=1}^{n}dP(x_i))P(X_{n+1}>s-(\sum_{j=1}^{n}x_j)\geq 0)} && (37)\\
+&\frac{\int_{x_1+\dots x_n\geq s,x_{n+2}+\dots x_{n+l}>t}(\Pi_{i=1,\neq n+1}^{n+l}dP(x_i))P(X_{n+1}>s-(\sum_{j=1}^{n}x_j)\geq 0)}{\int_{x_1+\dots x_n\geq s}(\Pi_{i=1}^{n}dP(x_i))P(X_{n+1}>s-(\sum_{j=1}^{n}x_j)\geq 0)} && (38)\\
=&P(X_{n+2}+\dots X_{n+l}\leq t,X_{n+1}+\dots X_{n+l}>t)+P(X_{n+2}+\dots X_{n+l}>t) && (39)
\end{aligned}
$$

これは、独立同一な指数分布 l 個についての確率であるから、n にはよらない。したがって、$N(t+s)-N(s),N(s)$ は独立である。補集合をとって計算をすすめると、

$$
\begin{aligned}
&P(N(t+s)-N(s)\geq l|N(s)=n)=P(N(t+s)-N(s)\geq l) && (40)\\
=&P((X_{n+2}+\dots X_{n+l}>t\vee X_{n+1}+\dots X_{n+l}\leq t)\wedge X_{n+2}+\dots X_{n+l}\leq t) && (41)\\
=&P(X_{n+1}+\dots X_{n+l}\leq t) && (42)\\
=&\int_0^t\lambda\exp(-\lambda s)\frac{(\lambda s)^{n-1}}{(n-1)!}ds && (43)\\
=&1-\exp(-\lambda t)\left(1+\dots\frac{(\lambda t)^{l-1}}{(l-1)!}\right) && (44)
\end{aligned}
$$

これはポアソン分布での $l,l+1,\dots\infty$ の事象確率にほかならない。したがって、増分はポアソン分布に従っている。

$t>s>u$ とすると、$N(t)-N(u),N(u)$ が独立で $N(s)-N(u),N(u)$ が独立であるから、$N(t)-N(u)=(N(t)-N(u))-(N(s)-N(u)),N(s)$ もまた独立である。同様に、時間区間が重ならなければすべて独立になる。

Proposition 2.4. ポアソン過程であれば、次を満たす。逆に以下を満たすような確率変数 $\{N(t)|t\geq 0\}$ があれば、それはポアソン過程と等しい。[*1]

1. $N(0)=0a.s.$
2. 独立増分であって、増分は $dt\lambda$ のポアソン分布に従う

Proof. この過程についての停止時間 $\sigma_n=\inf\{t|N(t)\geq n\}$ を定める。そして、この差分を $t_n=\sigma_n-\sigma_{n-1}$ でとる。これが独立でパラメータ λ についての指数分布であるとき、$\widehat{N}(t)=\max\{n|\sigma_n\leq t\}$ は明らかに $N(t)$ に一致して、しかも最初の構成にも一致する。よって、この停止時間の独立指数分布性を示せば十分である。ところが、停止時間は連続時間確率変数なので、条件付き確率を直接計算しようとすると都合が悪い。そこで、わずかに幅をもたせた条件付き確率を計算する。まず、t_2,t_1 について、t_1 は明らかに指数分布に従

[*1] ここでの等しいとは、$\{N(t)|t\geq 0\}$ の任意個を取り出した結合分布が同一であることを指すとする。

う。t_2 について、独立増分性を使うことで、

$$
\begin{aligned}
&P(t_2 > t | t_1 \in (s-\epsilon, s+\epsilon]) && (45)\\
\leq &\frac{P(N(t+s-\epsilon) - N(s+\epsilon) = 0, t_1 \in (s-\epsilon, s+\epsilon])}{P(t_1 \in (s-\epsilon, s+\epsilon])} && (46)\\
= &P(N(t+s-\epsilon) - N(s+\epsilon) = 0) && (47)\\
= &\exp(-\lambda(t-2\epsilon)) && (48)\\
&P(t_2 > t | t_1 \in (s-\epsilon, s+\epsilon]) && (49)\\
\geq &\frac{P(N(t+s+\epsilon) - N(s+\epsilon) = 0, t_1 \in (s-\epsilon, s+\epsilon])}{P(t_1 \in (s-\epsilon, s+\epsilon])} && (50)\\
= &\exp(-\lambda t) && (51)
\end{aligned}
$$

と、両側から評価できる。このとき、$A \in t_1^{-1}(\mathcal{B})$ について、A を被覆する $t_1^{-1}((s-\epsilon, s+\epsilon])$ の形をした集合の合併を B、A に含まれる、互いに交差しない $t_1^{-1}((s-\epsilon, s+\epsilon])$ の形をした集合の合併を C とする。半開区間はボレル族の生成元であるから、A と B, C の差は、$\epsilon \to 0$ に従って任意に小さくできる。このとき

$$
\begin{aligned}
&\int_A P(t_2 > t | t_1^{-1}(\mathcal{B}) dP && (52)\\
\leq &\int_B P(t_2 > t | t_1^{-1}(\mathcal{B}) dP && (53)\\
\leq &\int_B \exp(-\lambda(t-2\epsilon)) dP && (54)\\
&\int_A P(t_2 > t | t_1^{-1}(\mathcal{B}) dP && (55)\\
\geq &\int_C P(t_2 > t | t_1^{-1}(\mathcal{B}) dP && (56)\\
\geq &\int_C \exp(-\lambda t) dP && (57)
\end{aligned}
$$

ここで $\epsilon \to 0$ ととれば、$\exp(-\lambda t)$ は、$t_1^{-1}(\mathcal{B})$ 条件付き確率の資格を持つ。それは条件付けによらないため、t_1 と t_2 は独立である。

同様に、t_3, t_2, t_1 について議論すれば、t_3 が t_2, t_1 と独立な指数分布になることがわかる。すなわち、帰納的に $t_{n-1}, \ldots t_1$ の（微小区間での）条件下で、t_n が独立指数であることがわかる。

3 何が問題だったか

応用向けの書物では、例えば連続確率変数についての条件つき確率

$$
p(x = a | y = b) \tag{58}
$$

といった処理が度々なされる。ところで、この確率空間が本当に確率変数 x, y の像域に押し出されているのならば。この記法には曖昧さはない。これに意味をつけるには、x, y の同時確率密度分布 $p(x = a, y = b)$ を考えて、$p(y = b)$ で割ればよい。このときは、押し出し確率分布の絶対連続性だけが問題になる。ところが、$p(-|y = b)$ において、この第一引数を、測度論的な意味での測度と思いたい場合はどうなるだろうか？　つまり、$p(-|y = b)$ の第一引数は、加法族を受け入れる。したがって、集合演算や論理演算が有効であってほしい。

測度論的確率論では、条件付き確率というのは、条件付き期待値の特別な場合であり、それは、条件つけ変数の引き戻し加法族上の可測関数 $P(B|y)$ で、引き戻し加法族の可測集合 A について

$$\int_A P(B|y)dP = P(B \wedge A) \tag{59}$$

を満たすものである。

まず、この $P(B|y)$ は、B ごとに定義されるので、B を引数と思った時に測度にならないことがあるが、これについてはなんらかの修正ができたとしよう。しかしまだだめである。これは y の引き戻し加法族上の可測関数であって、y の像域上の可測関数ではない。しかしこれもまだ解消できる。y 引き戻し加法族上の可測関数は、y の像域を経由するような可測関数の合成に分解できる。すなわち、「一応」$P(B|y)$ は、y 像域上の測度関数と思うことができる。

それで、何が問題だったか？ $P(B|y)$ の B 側の引数について、$y = a$ という条件付による一点集合の逆像 $y^{-1}(\{a\})$ に、B を勝手に制限してよいのか？ということである。これは直感的にはできてほしい。例えば、y が連続変数ではないなら、条件付き確率は

$$P(B|y^{-1}(A)) = \frac{P(B \wedge y^{-1}(A))}{P(y^{-1}(A))} = P(B \wedge y^{-1}(A)|y^{-1}(A)) \tag{60}$$

となるので、この測度は条件付き集合上でしか値を取らないといっていい。しかし連続変数の時に、これができる、という保証がない。なぜなら、このときは $P(y^{-1}(A)) = 0$ になってしまうからだ。測度論的確率論の入門書にはこのような記述が見当たらなかったが、応用数学向けのポアソン過程の停止時間間隔の独立指数の証明は、殆どが、$\sigma_1 = t$ という連続変数の条件付け下で、条件付き確率の第一引数について、集合論的証明*2 をやっているものがほとんどだった。これはすくなくとも筆者にはギャップに見え、長時間拘泥してしまった。連続変数の像域値による条件付けと、その条件付き確率の測度論的演算が両立するのかどうかは未だにわからないが、ひとまずポアソン過程については、以上のように、若干幅をもたせれば、そのような一般論を要せず証明ができるはずである。

参考文献

[1] R. デュレット, 確率過程の基礎　丸善　(2012)

*2 以上での証明で処理したように、$t_2 > t_1$ という命題を、$N(-)$ の、しかも条件付け変数 s に依存した形の命題に変形できるのは、これらがともに、背後にある可測空間上の関数であることに基づいている。したがって、$p(x = a|y = b)$ のような有限語の確率変数の同時密度関数があればよい、というわけにはいかない。