確率論 IV

石綿 元

第 5 回講義

目次

1　分布関数・母関数及び中心極限定理

1.1　累積分布関数

「第 1 回講義　**データの視覚化**」において、度数と相対累積度数のヒストグラムを作成した。このとき、相対累積度数の持つ意味は、ある度数における割合を求めたり、ある割合以下の度数を知るためのグラフとして機能していた。確率分布は経験分布である度数分布の理論分布であった。度数分布と確率分布を対比して考えると、累積相対度数について対応するものが累積分布関数である。

すなわち、確率変数 X について

$$\mathcal{F}(t) = P(X \leq t)$$

である関数 $\mathcal{F}(t)$ を X の累積分布関数もしくは分布関数と定義する。

■離散型

$$\mathcal{F}(t) = P(X \leq t) = \sum_{x_i \leq t} f(x_i)$$

■連続型

$$\mathcal{F}(t) = P(X \leq t) = \int_{-\infty}^{t} f(x)dx$$

■確率密度関数と分布関数の関係　微分積分学の基本定理より、$\mathcal{F}'(x) = f(x)$。つまり、分布関数の接線の傾きが確率密度関数のその点での値 (高さ) であり、分布関数の値 (高さ) は、確率密度関数のその点までの面積である。

1.2　積率母関数

1.2.1　積率母関数（離散型）

確率変数 X の確率分布 $f(x_i)$ に対して、

$$M_X(t) = \sum_{i=1}^{\infty} e^{tx_i} f(x_i) = E[e^{tX}]$$

を離散型確率変数 X の積率母関数と定義する。

■ゼロ回微分（微分しないまま）　まず、定義 $M_X(t)$ に、そのまま $t = 0$ を代入すると、

$$M_X(0) = \sum_{i=1}^{\infty} e^0 f(x_i) = \sum_{i=1}^{\infty} f(x_i) = 1$$

である。全確率を表す。

■一回微分

$$M'_X(t) = \frac{d}{dt}M_X(t) = \sum_{i=1}^{\infty} x_i e^{tx_i} f(x_i)$$

ここで、$t = 0$ を代入すると、

$$M'_X(0) = \sum_{i=1}^{\infty} x_i f(x_i)$$

これは、原点周りの 1 次積率であり、つまりは期待値である。

■二回微分

$$M''_X(t) = \frac{d}{dt}M'_X(t) = \sum_{i=1}^{\infty} x_i^2 e^{tx_i} f(x_i)$$

ここで、$t = 0$ を代入すると、

$$M''_X(0) = \sum_{i=1}^{\infty} x_i^2 f(x_i)$$

これは、原点周りの 2 次積率 $E[X^2]$ になっている。したがって、分散と期待値の関係を用いて

$$V[X] = E[X^2] - (E[X])^2 = M''_X(0) - (M'_X(0))^2$$

が導出できる。

1.2.2 積率母関数（連続型）

確率変数 X の確率分布 $f(x)$ に対して、

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x)dx = E[e^{tX}]$$

を連続型確率変数 X の積率母関数と定義する。

■ゼロ回微分（微分しないまま）　まず、定義 $M_X(t)$ に、そのまま $t = 0$ を代入すると、

$$M_X(0) = \int_{-\infty}^{\infty} e^0 f(x)dx = \int_{-\infty}^{\infty} f(x)dx = 1$$

である。全確率を表す。

■一回微分

$$M'_X(t) = \frac{d}{dt}M_X(t) = \int_{-\infty}^{\infty} xe^{tx} f(x)dx$$

ここで、$t = 0$ を代入すると、

$$M'_X(0) = \int_{-\infty}^{\infty} xf(x)dx$$

これは、原点周りの 1 次積率であり、つまりは期待値である。

■二回微分

$$M''_X(t) = \frac{d}{dt}M'_X(t) = \int_{-\infty}^{\infty} x^2 e^{tx} f(x)dx$$

ここで、$t = 0$ を代入すると、

$$M''_X(0) = \int_{-\infty}^{\infty} x^2 f(x)dx$$

これは、原点周りの 2 次積率 $E[X^2]$ になっている。したがって、分散と期待値の関係を用いて

$$V[X] = E[X^2] - (E[X])^2 = M''_X(0) - (M'_X(0))^2$$

が導出できる。

1.3　確率母関数

1.3.1　確率母関数（離散型）

離散型確率変数 X の確率分布 $f(x_i)$ に対して、

$$G(z) = \sum_{i=1}^{\infty} z^{x_i} f(x_i) = E[z^X]$$

を確率変数 X の確率母関数と定義する。ここで、$z = e^t$ のとき、$M_X(t)$ である。

■全体確率　$z = 1$ を代入すると全確率が示される。

$$G(1) = \sum_{i=1}^{\infty} 1^{x_i} f(x_i) = \sum_{i=1}^{\infty} f(x_i) = 1$$

■期待値　$G(z)$ を z で一回微分する。

$$G'(z) = \frac{d}{dz}G(z) = \sum_{i=1}^{\infty} x_i z^{x_i - 1} f(x_i)$$

これに $z = 1$ を代入すると

$$G'(1) = \sum_{x=1}^{\infty} x_i 1^{x_i - 1} f(x_i) = \sum_{x=1}^{\infty} x_i f(x_i)$$

これは期待値である。

■分散　$G(z)$ を z で二回微分する。

$$G''(z) = \frac{d^2}{dz^2}G(z) = \sum_{i=1}^{\infty} x_i(x_i - 1) z^{x_i - 2} f(x_i)$$

これに $z = 1$ を代入すると

$$\begin{aligned} G''(1) &= \sum_{i=1}^{\infty} x_i(x_i - 1) 1^{x_i - 2} f(x_i) \\ &= \sum_{i=1}^{\infty} x_i^2 f(x_i) - \sum_{i=1}^{\infty} x_i f(x_i) \end{aligned}$$

ここで、右辺第 1 項目は $E[X^2]$ である。よって、

$$E[X^2] = G''(1) + G'(1)$$

分散と期待値の関係から

$$V[X] = E[X^2] - (E[X])^2 = G''(1) + G'(1) - \{G'(1)\}^2$$

とすることで、分散が求まる。

1.3.2　確率母関数（連続型）

連続型確率変数 X の確率分布 $f(x)$ に対して、

$$G(z) = \int_{-\infty}^{\infty} z^x f(x)dx = E[z^X]$$

を確率変数 X の確率母関数と定義する。ここで、$z = e^t$ のとき、$M_X(t)$ である。

■全体確率　$z = 1$ を代入すると全確率が示される。

$$G(1) = \int_{-\infty}^{\infty} 1^x f(x)dx = \int_{-\infty}^{\infty} f(x)dx = 1$$

■期待値　$G(z)$ を z で一回微分する。

$$G'(z) = \frac{d}{dz}G(z) = \int_{-\infty}^{\infty} xz^{x-1} f(x)dx$$

これに $z = 1$ を代入すると

$$G'(1) = \int_{-\infty}^{\infty} x1^{x-1} f(x)dx = \int_{-\infty}^{\infty} xf(x)dx$$

これは期待値である。

■分散　$G(z)$ を z で二回微分する。

$$G''(z) = \frac{d^2}{dz^2}G(z) = \int_{-\infty}^{\infty} x(x-1)z^{x-2} f(x)dx$$

これに $z = 1$ を代入すると

$$\begin{aligned} G''(1) &= \int_{-\infty}^{\infty} x(x-1)1^{x-2} f(x)dx \\ &= \int_{-\infty}^{\infty} x^2 f(x)dx - \int_{-\infty}^{\infty} xf(x)dx \end{aligned}$$

ここで、右辺第 1 項目は $E[X^2]$ である。よって、

$$E[X^2] = G''(1) + G'(1)$$

分散と期待値の関係から

$$V[X] = E[X^2] - (E[X])^2 = G''(1) + G'(1) - \{G'(1)\}^2$$

とすることで、分散が求まる。

1.4 中心極限定理

1.4.1 チェビシェフの不等式

確率変数 X について、平均 $\mu = E[X]$、分散 $\sigma^2 = V[X]$ とするとき、任意の $\varepsilon > 0$ に対して

$$P(|X-\mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

が成立する。

■Proof 分散の定義から確認してみる。

$$V[X] = \sigma^2 = \int_{-\infty}^{\infty}(x-\mu)^2 f(x)dx = \int_{-\infty}^{\mu-\varepsilon}(x-\mu)^2 f(x)dx + \int_{\mu-\varepsilon}^{\mu+\varepsilon}(x-\mu)^2 f(x)dx + \int_{\mu+\varepsilon}^{\infty}(x-\mu)^2 f(x)dx$$

$$\geq \int_{-\infty}^{\mu-\varepsilon}(x-\mu)^2 f(x)dx + \int_{\mu+\varepsilon}^{\infty}(x-\mu)^2 f(x)dx$$

ここで、x は、任意 ε よりも外側を動くので、$x-\mu$ よりも ε のほうが小さいから、

$$\geq \int_{-\infty}^{\mu-\varepsilon}\varepsilon^2 f(x)dx + \int_{\mu+\varepsilon}^{\infty}\varepsilon^2 f(x)dx = \varepsilon^2\int_{-\infty}^{\mu-\varepsilon} f(x)dx + \varepsilon^2\int_{\mu+\varepsilon}^{\infty} f(x)dx = \varepsilon^2 P(|X-\mu| \geq \varepsilon)$$

よって、

$$\sigma^2 \geq \varepsilon^2 P(|X-\mu| \geq \varepsilon)$$

1.4.2 大数の法則

確率変数 X が、平均 $\mu = E[X]$、分散 $\sigma^2 = V[X]$ のとある確率分布に従うとする。

いま、n 個のまったく独立な確率変数 $X_1,\ X_2,\ X_3,\ \cdots X_n$ が、すべて独立に確率変数 X と同一の確率分布に従うとする。

この同一の確率分布に従う n 個のまったく独立な確率変数 $X_1,\ X_2,\ X_3,\ \cdots X_n$ の総平均を

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \cdots + X_n}{n}$$

とおく。

このとき、n が十分大きいならば、$\bar{X}$ が $\mu = E[X]$ に近づく。すなわち、

$$\lim_{n\to\infty} P(|\bar{X}-\mu| < \varepsilon) = 1$$

■理論的確認 総平均の期待値を $E[\bar{X}] = \mu$ また、総平均の分散は、$V[\bar{X}] = \frac{\sigma^2}{n}$ であるとして代入すると、チェビシェフの不等式から、

$$P(|\bar{X}-\mu| \geq \varepsilon) = P(|\bar{X}-E[\bar{X}]| \geq \varepsilon) \leq \frac{V[\bar{X}]}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}$$

このとき、

$$\lim_{n\to\infty}\frac{\sigma^2}{n\varepsilon^2} = 0$$

つまり、両端がゼロに近づく確率分布の総面積となるので、

$$\lim_{n\to\infty} P(|\bar{X}-\mu| < \varepsilon) = 1$$

1.4.3　中心極限定理

■**中心極限定理-[[本節は非常に重要である]]**　確率変数 X が、平均 $\mu = E[X]$、分散 $\sigma^2 = V[X]$ のとある確率分布に従うとする。

いま、n 個のまったく独立な確率変数 $X_1,\ X_2,\ X_3,\ \cdots X_n$ が、すべて独立に確率変数 X と同一の確率分布に従うとする。

この同一の確率分布に従う n 個のまったく独立な確率変数 $X_1,\ X_2,\ X_3,\ \cdots X_n$ の総平均を

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \cdots + X_n}{n}$$

とおく。

このとき、n が十分大きいならば、**確率変数 X が従う分布によらず、**総和の平均 $\bar{X}$ は平均 μ、分散 $\frac{\sigma^2}{n}$ の**正規分布 $N\left(\mu,\ \frac{\sigma^2}{n}\right)$ に近づく。**

さらに、確率変数 X が「平均 μ、分散 σ^2 のある分布」に従うとき、$\bar{X}$ ではなく、X と同一の確率分布に従う n 個のまったく独立な確率変数 X_i の和を確率変数 S とすると、

$$S = X_1 + X_2 + X_3 + \cdots + X_n = \sum_{i=1}^{n} X_i$$

も中心極限定理によって正規分布に近づくが、**その場合の正規分布は $\mathbf{N}(n\mu, n\sigma^2)$ に近づく。**

■**正規近似**　中心極限定理より、同一の分布に従う n 個の確率変数の総和の平均は n が大きいとき正規分布に従うと考えてよい。これには、離散型、連続型かについても関係ない。精密さを要求しない場合、おおむね **30 個以上**程度でも正規近似できると考えてもよい。

特に、２項分布の試行回数 n が大きく、かつ、p が小さくないとき、正規分布に近似される。

例さいころを３０回転がす試行について１の目が出るかどうかに着目すると、ベルヌーイ試行となり、確率変数 X は、２項分布に従う。したがって、$\mathrm{B}(30, \frac{1}{6})$ である。このとき、この２項分布の平均は、$np = 30 \times \frac{1}{6} = 5$ であり、分散は $np(1-p) = 30 \times \frac{1}{6} \times \left(1 - \frac{1}{6}\right) = 5 \times \frac{5}{6} = \frac{25}{6}$ である。

$$f(x) = {}_{30}C_x \left(\frac{1}{6}\right)^x \left(1 - \frac{1}{6}\right)^{30-x}$$

この平均と分散を持つ正規分布は、$\mathrm{N}(5, \frac{25}{6})$

$$f(x) = \frac{1}{\sqrt{2\pi \frac{25}{6}}} e^{-\frac{1}{2}\frac{(x-5)^2}{\frac{25}{6}}}$$

さいころを１２０回転がす試行について、再び、１の目が出るかどうかに着目するベルヌーイ試行の確率変数 X は、２項分布 $\mathrm{B}(120, \frac{1}{6})$ である。このとき、この２項分布の平均は、$np = 120 \times \frac{1}{6} = 20$ であり、分散は $np(1-p) = 120 \times \frac{1}{6} \times \left(1 - \frac{1}{6}\right) = 20 \times \frac{5}{6} = \frac{100}{6}$ である。

$$f(x) = {}_{120}C_x \left(\frac{1}{6}\right)^x \left(1 - \frac{1}{6}\right)^{120-x}$$

この平均と分散を持つ正規分布は、$\mathrm{N}(20, \frac{100}{6})$

$$f(x) = \frac{1}{\sqrt{2\pi \frac{100}{6}}} e^{-\frac{1}{2}\frac{(x-20)^2}{\frac{100}{6}}}$$

2　確率過程 < 参考 >

2.1　確率過程

時間 t をパラメータに持つ確率変数の族 X_t を確率過程と定義する。t が決まれば、X_t は 1 つの確率変数となる。

2.2　マルコフ連鎖

2.2.1　マルコフ連鎖の定義

状態空間 $S=\{1,2,3,\cdots,N\}$ 上の離散時間確率過程である確率変数列 $\{X_n;n=0,1,2,\cdots\}$ が任意の時刻列 $n=1,2,\cdots$ 任意の状態列 $s_0,s_1,\cdots,s_n\in S$ に対して、次の条件付き確率

$$P\left(\{X_n(\omega)=s_n\}|\{X_0(\omega)=s_0,\cdots,X_{n-1}(\omega)=s_{n-1}\}\right)=P\left(\{X_n(\omega)=s_n\}|\{X_{n-1}(\omega)=s_{n-1}\}\right)$$

が成立するとき、$\{X_n\}$ は、S 上のマルコフ連鎖と定義する。

$X_0\sim X_{n-1}$ が起こった条件の下で X_n が起こる条件付き確率が X_n の直前である現在の状態 X_{n-1} のみによって決まり、$X_0\sim X_{n-2}$ とは独立である。この性質をマルコフ性と呼ぶ。

■推移確率　S 上のマルコフ連鎖 $\{X_n;n=0,1,2,\cdots\}$ に対して、

$$P\left(\{X_n(\omega)=j\}|\{X_{n-1}(\omega)=i\}\right)=P_{ij}^{(n)}$$

とおくとき、$P_{ij}^{(n)}$ を推移確率と定義する。

さらに、$P_{ij}^{(n)}$ を (i,j) 成分とする N 次正方行列

$$P^{(n)}=\left[P_{ij}^{(n)}\right]=\begin{bmatrix} p(1\to1) & p(1\to2) & \cdots & p(1\to N)\\ p(2\to1) & p(2\to2) & \cdots & p(2\to N)\\ \cdots & \cdots & \cdots & \cdots\\ p(N\to1) & p(N\to2) & \cdots & p(N\to N)\end{bmatrix}_{i\to j}$$

この行列を時刻 $n-1$ におけるマルコフ連鎖 $\{X_n\}$ の推移行列と定義する。

2.2.2　チャップマン・コロモゴロフの方程式

$\forall m,\forall n\in\boldsymbol{N}$ のとき、

$$P^{(m+n)}=P^{(m)}P^{(n)}$$

が成立する。このとき、推移確率は $P^{(m+n)}$ の各成分である。

2.2.3　マルコフ連鎖の例と状態推移図

例 1）　状態 $S=\{0,1\}$、時刻 $n-1$ で

$$\begin{array}{lll} 0\to0 & \text{に変わる確率} & 1-\varepsilon_n\\ 0\to1 & \text{に変わる確率} & \varepsilon_n\\ 1\to0 & \text{に変わる確率} & p_n\\ 1\to1 & \text{に変わる確率} & 1-p_n \end{array}$$

とするとき、

$$
\begin{aligned}
P_{00}^{(n)} &= 1-\varepsilon_n \\
P_{01}^{(n)} &= \varepsilon_n \\
P_{10}^{(n)} &= p_n \\
P_{11}^{(n)} &= 1-p_n
\end{aligned}
$$

推移行列は、

$$P^{(n)} = \begin{bmatrix} 1-\varepsilon_n & \varepsilon_n \\ p_n & 1-p_n \end{bmatrix}$$

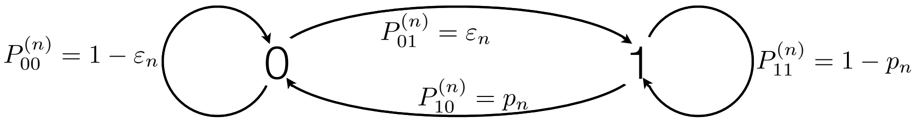

図 1　状態推移図

2.3　ポアソン過程と Γ-分布

2.3.1　ポアソン過程の定義

■定義　事象が時刻 $t=0$ から t までに発生した数を $N(t)$ とする。この $N(t)$ が実現する確率について、

1. $P(N(0)=0)=1$
2. 十分小さい時間幅 h について $P(N(t+h)-N(t)=1)=\lambda h+o(h) \quad (h\to 0)$
3. 十分小さい時間幅 h について $P(N(t+h)-N(t)\geq 2)=o(h) \quad (h\to 0)$
4. 時刻 $t_0<t_1<t_2<\cdots<t_n$ に対して、確率変数 $N(t)$ は、各幅 t_i-t_{i-1} の幅のみに依存（これを定常増分という）して、互いに排反な区間の増分 $N(t_2)-N(t_1), N(t_3)-N(t_2),\cdots,N(t_n)-N(t_{n-1})$ は、互いに独立（これを独立増分という)。（このとき $N(t)$ は定常独立増分であるという）

この 4 つを満たす $N(t)$ をポアソン過程と定義する。ここで、パラメータ λ は正数。$o(h)$ はランダウ記号[*1]。

■定理　ポアソン過程 $N(t)$ は、平均 $\mu=\lambda t$ のポアソン分布である。

Proof) $N(t)$ の具体的な形を考えていく。いま、$N(t)=k$ について、とりうる確率をすべて表現すると、$k=0,1,2,3,\cdots$ として、すべての確率の和は、当然、

$$P_0(t)+P_1(t)+P_2(t)+P_3(t)+\cdots=1$$

これら各 $P_k(t)$ に共通な一般化した形が確率変数 $N(t)$ をもつ関数である。まず、$P_0(t)$ から順に考えていくと、定義の条件 2. と 3. よりすぐに

$$P_1(h)\approx\lambda h,\quad P_2(h)+P_3(h)+\cdots\approx 0$$

[*1] ランダウ記号の定義：

$$\lim_{x\to a}\frac{f(x)}{h(x)}=0 \Rightarrow f(x)=o(h(x)) \quad (x\to a)$$

よって、

$$P_0(h) = 1 - (P_1(h) + P_2(h) + P_3(h) + \cdots) = 1 - \lambda h$$

いま、h の幅とは $(t, t+h]$ であって、$(0, t]$ とは排反であるから定義の条件 4. より、独立事象の確率であるから、

$$P_0(t+h) = P_0(t) \cdot P_0(h) \approx P_0(t)(1 - \lambda h) = P_0(t) - \lambda h P_0(t)$$

が成り立つ。よって、

$$\begin{aligned} \frac{P_o(t+h) - P_0(t)}{h} &= -\lambda P_0(t) \\ \lim_{h \to 0} \frac{P_o(t+h) - P_0(t)}{h} &= \lim_{h \to 0} \{-\lambda P_0(t)\} \\ P_0'(t) &= -\lambda P_0(t) \end{aligned}$$

ここで、変数分離形として、

$$\begin{aligned} \frac{dP_0(t)}{dt} &= -\lambda P_0(t) \\ \int \frac{1}{P_0(t)} dt &= \int -\lambda dt + \gamma \\ \log P_0(t) &= -\lambda t + \gamma \\ P_0(t) &= e^{-\lambda t + \gamma} = e^{\gamma} e^{-\lambda t} \end{aligned}$$

ここで、$e^{\gamma} = C$ とおくと、$P_0(t) = Ce^{-\lambda t}$。また、定義の条件 1. より、$P_0(0) = 1$ だから、$P_0(0) = Ce^{-\lambda 0} = 1$ より $C = 1$。つまり、

$$P_0(t) = e^{-\lambda t}$$

次に、$k = 1$ のときを考えると、$(0, t]$ で 1 回発生もしくは $(t, t+h]$ で 1 回発生であるから、

$$P_1(t+h) = P_1(t) \cdot P_0(h) + P_0(t) \cdot P_1(h) = P_1(t) \cdot (1 - \lambda h) + P_0(t) \cdot (\lambda h)$$

よって、

$$\begin{aligned} \frac{P_1(t+h) - P_1(t)}{h} &= -\lambda P_1(t) + \lambda P_0(t) \\ \lim_{h \to 0} \frac{P_1(t+h) - P_1(t)}{h} &= \lim_{h \to 0} \{-\lambda P_1(t) + \lambda P_0(t)\} \\ P_1'(t) &= -\lambda P_1(t) + \lambda P_0(t) \\ P_1'(t) + \lambda P_1(t) &= \lambda P_0(t) \end{aligned}$$

線形微分方程式なので、変数 t の取り得る範囲は 0 から t までだから、

$$\begin{aligned} P_1(t) &= e^{-\int_0^t \lambda ds} \left\{ \int_0^t \lambda P_0(s) e^{\int_0^t \lambda ds} ds + \gamma \right\} \\ &= e^{-\lambda t} \left\{ \int_0^t \lambda P_0(s) e^{\lambda s} ds + \gamma \right\} \\ &= e^{-\lambda t} \int_0^t \lambda e^{\lambda s} P_0(s) ds + e^{-\lambda t} \gamma \end{aligned}$$

ここで、$e^{-\lambda t}\gamma = C$ とおくと、

$$P_1(t) = e^{-\lambda t}\int_0^t \lambda e^{\lambda s} P_0(s)ds + C$$

いま、$t=0$ を代入してみると、積分区間は 0 から 0 までの積分で、１点での積分はゼロだから、$P_1(0) = C$ とわかる。よって、

$$P_1(t) = e^{-\lambda t}\int_0^t \lambda e^{\lambda s} P_0(s)ds + P_1(0) \tag{1}$$

時刻 0 での発生回数はないので、発生回数がない確率は $P_0(0) = 1$ であるが、発生回数が 1 の場合の発生確率 $P_1(0) = 0$ である。

全く同様に k 回の場合を考えると、$(0, t]$ での発生回数と $(t, t+h]$ での発生回数をあわせて k であるから、

$$P_k(t+h) = P_k(t)\cdot P_0(h) + P_{k-1}(t)\cdot P_1(h) + P_{k-2}(t)\cdot P_2(h) + \cdots$$

いま、定義の条件 3. より、$P_2(h), P_3(h), \cdots, \approx 0$ だから、

$$P_k(t+h) \approx P_k(t)\cdot(1-\lambda h) + P_{k-1}(t)\cdot(\lambda h)$$

よって、$k=1$ の場合の式 (1) を導出したときと全く同様に

$$\begin{aligned}\frac{P_k(t+h)-P_k(t)}{h} &= -\lambda P_k(t) + \lambda P_{k-1}(t)\\ P'_k(t) &= -\lambda P_k(t) + \lambda P_{k-1}(t)\\ P_k(t) &= e^{-\lambda t}\int_0^t \lambda e^{\lambda s} P_{k-1}(s)ds + P_k(0)\end{aligned}$$

時刻 0 での発生回数はないので、発生回数が 1 以上の場合の発生確率 $P_k(0) = 0 \quad (k \neq 0)$ である。よって、

$$P_k(t) = e^{-\lambda t}\int_0^t \lambda e^{\lambda s} P_{k-1}(s)ds \tag{2}$$

式 (2) を具体的に解いていく。

$$\begin{aligned}P_1(t) &= e^{-\lambda t}\int_0^t \lambda e^{\lambda s} P_0(s)ds = e^{-\lambda t}\int_0^t \lambda e^{\lambda s} e^{-\lambda s} ds = e^{-\lambda t}\lambda[s]_0^t = \lambda t e^{-\lambda t}\\ P_2(t) &= e^{-\lambda t}\int_0^t \lambda e^{\lambda t} P_1(s)ds = e^{-\lambda t}\int_0^t \lambda e^{\lambda s}\lambda s e^{-\lambda s} ds = e^{-\lambda t}\int_0^t \lambda^2 s ds = e^{-\lambda t}\lambda^2[\frac{1}{2}s^2]_0^t = e^{-\lambda t}\frac{(\lambda t)^2}{2!}\\ P_3(t) &= e^{-\lambda t}\int_0^t \lambda e^{\lambda s} e^{-\lambda s}\frac{(\lambda s)^2}{2!}ds = e^{-\lambda t}\lambda^3\frac{1}{2!}[\frac{1}{3}s^3]_0^t = e^{-\lambda t}\frac{(\lambda t)^3}{3!}\\ &\vdots\\ P_k(t) &= e^{-\lambda t}\frac{(\lambda t)^k}{k!}\end{aligned}$$

つまり、ポアソン過程 $N(t)$ は、平均 $\mu = \lambda t$ のポアソン分布である。

2.3.2 Γ-分布の導出

連続型確率密度関数とその性質に関しては、前回までの講義で解説したが、理解のしやすい連続型一様分布を除くと、指数分布と χ^2-分布については、Γ-分布の特別な場合であり、F-分布は、χ^2-分布どうしの組み合わせで、t-分布は、標準正規分布と χ^2-分布の組み合わせで導出できた。では、正規分布と Γ-分布どのように導出されるのだろうか。正規分布については、次の講義回で導出と統計上の意味・意義を詳しく解説する。ここでは、Γ-分布について調べてみよう。

前節のポアソン過程において、α 以上の値をとる確率は、分布関数を考慮して、

$$P(X_t \geq \alpha) = 1 - \sum_{k=0}^{\alpha-1} e^{-\lambda t}\frac{(\lambda t)^k}{k!} = 1 - e^{-\lambda t}\sum_{k=0}^{\alpha-1}\frac{(\lambda t)^k}{k!} \equiv \mathcal{D}(t)$$

この $\mathcal{D}(t)$ について、時刻 t での微少変化を考える。(α 以上が起こる時間変化を求める)

$$\begin{aligned}\frac{d\mathcal{D}(t)}{dt} &= \lambda e^{-\lambda t}\sum_{k=0}^{\alpha-1}\frac{(\lambda t)^k}{k!} - e^{-\lambda t}\sum_{k=0}^{\alpha-1}\frac{k(\lambda t)^{k-1}\lambda}{k!} \\ &= \lambda e^{-\lambda t}\sum_{k=0}^{\alpha-1}\frac{(\lambda t)^k}{k!} - \lambda e^{-\lambda t}\sum_{k=0}^{\alpha-1}\frac{k(\lambda t)^{k-1}}{k!}\end{aligned}$$

ここで、$\sum$ の $k=0$ を具体的に代入すると、

$$\begin{aligned}&= \lambda e^{-\lambda t} + \lambda e^{-\lambda t}\sum_{k=1}^{\alpha-1}\frac{(\lambda t)^k}{k!} - \lambda e^{-\lambda t}\sum_{k=1}^{\alpha-1}\frac{k(\lambda t)^{k-1}}{k!} \\ &= \lambda e^{-\lambda t}\left\{1 + \sum_{k=1}^{\alpha-1}\frac{(\lambda t)^k}{k!} - \sum_{k=1}^{\alpha-1}\frac{k(\lambda t)^{k-1}}{k!}\right\} \\ &= \lambda e^{-\lambda t}\left[1 + \sum_{k=1}^{\alpha-1}\left\{\frac{(\lambda t)^k}{k!} - \frac{(\lambda t)^{k-1}}{(k-1)!}\right\}\right]\end{aligned}$$

ここで、$\sum$ の中を計算すると、$k=1$ と $k=\alpha-1$ のみ残り、残りの項はすべてカスケード状に消える。

$$\begin{aligned}&= \lambda e^{-\lambda t}\left[1 + \frac{(\lambda t)^1}{1!} - \frac{(\lambda t)^0}{0!} + \frac{(\lambda t)^2}{2!} - \frac{(\lambda t)^1}{1!} + \cdots + \frac{(\lambda t)^{\alpha-2}}{(\alpha-2)!} - \frac{(\lambda t)^{\alpha-3}}{(\alpha-3)!} + \frac{(\lambda t)^{\alpha-1}}{(\alpha-1)!} - \frac{(\lambda t)^{\alpha-2}}{(\alpha-2)!}\right] \\ &= \lambda e^{-\lambda t}\frac{(\lambda t)^{\alpha-1}}{(\alpha-1)!} \\ &= \frac{\lambda^\alpha}{(\alpha-1)!}e^{-\lambda t}t^{\alpha-1}\end{aligned}$$

ここで、$\lambda = \frac{1}{\beta}$ と置き、$(\alpha-1)! = \Gamma(\alpha)$ であるから、

$$= \frac{1}{\beta^\alpha\Gamma(\alpha)}t^{\alpha-1}e^{-\frac{t}{\beta}}$$

この確率変数 t は、時刻 t での微少変化量だから、x で表すとすると、

$$f(x) = \frac{1}{\beta^\alpha\Gamma(\alpha)}x^{\alpha-1}e^{-\frac{x}{\beta}}$$

Γ-分布は、単位時間あたり $\frac{1}{\beta}$ で起こる事象が α 回起こるまでの時間を表し、機械や電子部品の寿命モデルなどに現れてくる。

索引