1 preliminaries

1.1 probability

S : a sample space

$E \subset S$: an event

$P(E)$: the probability of E

Axiom (1) $0 \leq P(E) \leq 1$

Axiom (2) $P(S) = 1$

Axiom (3) $\{E_1, E_2, \ldots\}$, $E_i \cap E_j = \phi$ $(i \neq j)$ に対して

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i)$$

1. もし $E \subset F$ ならば $P(E) \leq P(F)$
2. $P(E^c) = 1 - P(E)$
3. $\{E_i\}$ が排反ならば $P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i)$
4. $P\left(\bigcup_{i=1}^{\infty} E_i\right) \leq \sum_{i=1}^{n} P(E_i)$

$\{E_n\}_{1 \leq n}$ が増加列であるとは $E_n \subset E_{n+1}$ $(1 \leq n)$ が成立するときにいい、減少列であるとは $E_n \supset E_{n+1}$ $(1 \leq n)$ が成立するときにいう。$\{E_n\}_{1 \leq n}$ が増加列であるとき

$$\lim_{n \to \infty} E_n = \bigcup_{i=1}^{\infty} E_i$$

$\{E_n\}_{1 \leq n}$ が減少列であるとき

$$\lim_{n \to \infty} E_n = \bigcap_{i=1}^{\infty} E_i$$

と定義する。

Proposition 1 もし $\{E_n\}_{n \geq 1}$ が増加列、または減少列であるとき、

$$\lim_{n \to \infty} P(E_n) = P\left(\lim_{n \to \infty} E_n\right)$$

が成立する。

Proof $\{E_i\}$ が増加列のときを示す。$F_1 = E_1, F_i = E_i \setminus E_{i-1}$ $(i > 1)$ とおけば、

$$\bigcup_{i=1}^{n} F_i = \bigcup_{i=1}^{n} E_i, \quad \bigcup_{i=1}^{\infty} F_i = \bigcup_{i=1}^{\infty} E_i,$$

であるから

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = P\left(\bigcup_{i=1}^{\infty} F_i\right) = \sum_{i=1}^{\infty} P(F_i) = \lim_{n\to\infty} \sum_{i=1}^{n} P(F_i) = \lim_{n\to\infty} P\left(\bigcup_{i=1}^{n} F_i\right)$$
$$= \lim_{n\to\infty} P\left(\bigcup_{i=1}^{n} E_i\right) = \lim_{n\to\infty} P(E_n)$$

減少列の場合には補集合を考える。

Proposition 2 The Borel-Cantelli Lemma
もし $\sum_{i=1}^{\infty} P(E_i) < \infty$ ならば、$P($無限の E_i が起きる$) = 0$ が成立する。

Proof 無限の E_i が起きる event とは

$$\limsup_{i\to\infty} E_i = \bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} E_i$$

のことである。proposition によって

$$P\left(\bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} E_i\right) = P\left(\lim_{n\to\infty} \bigcup_{i=n}^{\infty} E_i\right) = \lim_{n\to\infty} P\left(\bigcup_{i=n}^{\infty} E_i\right) \le \lim_{n\to\infty} \sum_{i=n}^{\infty} P(E_i) = 0$$

$\{E_i\}$ が独立であるとは、任意の部分列 $\{E_{i_j}\}_{1\le j\le m}$ に対して

$$P\left(\bigcap_{j=1}^{m} E_{i_j}\right) = \prod_{j=1}^{m} P(E_{i_j})$$

が成立するときにいう。この場合、$E'_i = E_i$ or E^c_i とするとき

$$P\left(\bigcap_{j=1}^{m} E'_{i_j}\right) = \prod_{j=1}^{m} P(E'_{i_j})$$

も成立する。

Proposition 3 Converse to the Borel-Cantelli Lemma
もし $\{E_i\}_{1\le i}$ が独立で、$\sum_{n=1}^{\infty} P(E_n) = 1$ ならば $P($無限の E_i が起きる$) = 1$ が成立する。

Proof

$$P(\text{無限の } E_i \text{ が起きる}) = P\left(\lim_{n\to\infty}\bigcup_{i=n}^{\infty} E_i\right) = \lim_{n\to\infty} P\left(\bigcup_{i=n}^{\infty} E_i\right) = \lim_{n\to\infty}\left(1 - P\left(\bigcap_{i=n}^{\infty} E_i\right)\right)$$

に

$$P\left(\bigcap_{i=n}^{\infty} E_i^c\right) = \prod_{i=n}^{\infty}(1 - P(E_i)) \leq \prod_{i=n}^{\infty} e^{-P(E_i)} = \exp\left(-\sum_{i=n}^{\infty} P(E_i)\right) = 0$$

を適用する。

1.2 確率変数

S から実数への関数 $X : S \to \mathbf{R}$ を確率変数という。$A \subset \mathbf{R}$ に対して

$$P\{X \in A\} = P(X^{-1}(A))$$

とおく。X の分布関数 $F(x)$ は

$$F(x) = P\{X \le x\} = P\{X \in (-\infty, x)\}$$

によって定義される。また、$\overline{F}(x) = 1 - F(x)$ とすれば、

$$\overline{F}(x) = P\{X > x\}$$

という意味がある。確率変数は可能な値が可算であるとき、離散的であるといわれる。したがって

$$F(x) = \sum_{y \le x} P\{X = y\}$$

また、

$$F(x) = \int_{-\infty}^{x} f(t)dt, \quad f(x) = \frac{dF(x)}{dx}$$

と表されるとき、$F(x)$ は連続的であるといわれる。このとき $f(x)$ は F の確率密度関数とよばれる。

X, Y を確率変数とする。

$$F(x, y) = P\{X \le x, Y \le y\}$$

を結合確率分布という。X, Y の確率分布を $F_X(x), F_Y(x)$ とすると

$$F_X(x) = \lim_{y \to \infty} F(x, y), \quad F_Y(x) = \lim_{x \to \infty} F(x, y)$$

が成立する。左は $\bigcup_{n=1}^{\infty}\{X \le x, Y \le y_n\} = \{X \le x\}$ からわかる。ただし、$\{y_n\}$ は $\lim_{n \to \infty} y_n = \infty$ なる任意の数列である。

$$F(x, y) = F_X(x)F_Y(y)$$

が成立するとき X, Y は独立であるといわれる。一般に、n この確率分布 $X_1, \ldots, X_n$ に対して結合確率分布

$$F(x_1, \ldots, x_n) = P\{X_1 \le x_1, \ldots, X_n \le x_n\}$$

が定義される。$X_1, \ldots, X_n$ が独立であるとは

$$F(x_1, \ldots, x_n) = F_{X_1}(x_1)F_{X_2}(x_2)\cdots F_{X_n}(x_n)$$

であるときにいう。ここに

$$F_{X_i}(x_i) = \lim_{\substack{x_j \to \infty \\ j \neq i}} F(x_1, \ldots, x_n)$$

$$F(x_1, \ldots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f(x_1, \ldots, x_n) dx_1 \cdots dx_n$$

と表されるとき、F は連続であるといわれる。

1.3 期待値

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x dF(x) \\ &= \begin{cases} \displaystyle\int_{-\infty}^{\infty} x f(x) dx & X \text{ が連続のとき} \\ \displaystyle\sum_x x P(X = x) & X \text{ が離散的のとき} \end{cases} \end{aligned}$$

が存在するとき、これを期待値という。h をほかの確率変数とすると、

$$E[h(X)] = \int_{-\infty}^{\infty} h(x) dF(x)$$

となる。
X の分散は

$$VarX = E[(X - EX)^2] = E[X^2] - E^2[X]$$

で定義され、X, Y の共分散は

$$Cov(X, Y) = E[(X - EX)(Y - EY)] = E[XY] - E[X]E[Y]$$

で与えられる。相関係数は

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sqrt{VarX}\sqrt{VarY}}$$

となる。

$$E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i]$$

$$Var\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} Var(X_i) + 2\sum_{i<j} Cov(X_i, X_j)$$

example　$A_1, \ldots, A_n$ を events とし、indicator 関数を

$$I_j = \begin{cases} 1 & A_j \text{ が起きる} \\ 0 & \text{そのほか} \end{cases}$$

とする。

$N = \sum_{j=1}^{n} I_j$ とおく。2 項公式

$$(a+b)^N = \sum_{i=0}^{N} \binom{N}{i} a^{N-i} b^i$$

より、$N \le n,\ \binom{N}{i} = 0\ (i > N)$ に注意して

$$(1-1)^N = \sum_{i=0}^{n} \binom{N}{i} (-1)^i$$

左辺は $N = 0$ のとき 1 と約束する。

$$I = \begin{cases} 1 & N > 0 \\ 0 & N = 0 \end{cases}$$

を導入すれば

$$I = \sum_{i=1}^{n} \binom{N}{i} (-1)^{i+1}$$

を得る。よって

$$E[I] = E[N] - E\left[\binom{N}{2}\right] + \cdots + (-1)^{n+1} E\left[\binom{N}{n}\right]$$

これは

$$\begin{aligned} P\left(\bigcup_{i=1}^{n} A_i\right) = & \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) \\ & + \cdots + (-1)^{n+1} P(A_1 \cap A_2 \cap \ldots A_n) \end{aligned}$$

を意味する。

events $A_1, \ldots, A_n$ のちょうど r こ起きるという event を考える。そのため

$$I_r = \begin{cases} 1 & N = r \\ 0 & N \neq r \end{cases}$$

とおく。

$$I_r = \binom{N}{r}(1-1)^{N-r} = \binom{N}{r}\sum_{i=0}^{n-r}\binom{N-r}{i}(-1)^i = \sum_{i=0}^{n-r}\binom{N}{r+i}\binom{r+i}{r}(-1)^i$$

より

$$E[I_r] = \sum_{i=0}^{n-r}(-1)^i\binom{r+i}{r}E\left[\binom{N}{r+i}\right]$$

すなわち

$$P(A_1,\ldots,A_n \text{ のちょうど } r \text{ こ起きる}) = \sum_{i=0}^{n-r}(-1)^i\binom{r+i}{r}\sum_{j_1<j_2<\cdots<j_{r+i}}P(A_{j_1}\cap A_{j_2}\cap\cdots\cap A_{j_{r+i}})$$

問題　N を非負の整数をとる確率変数とするとき

$$E[N] = \sum_{i=1}^{\infty}P\{N \geq i\}$$

であることを示せ。より一般に X が非負の確率変数で、F を分布とするとき

$$E[X] = \int_0^{\infty}\overline{F}(x)dx$$

であり、

$$E[X^n] = \int_0^{\infty}nx^{n-1}\overline{F}(x)dx$$

であることを証明せよ。

1.4 moment

確率変数 X のモーメント母関数を

$$\psi_X(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} dF(x)$$

によって定義する。以下では考察する積分の存在を仮定する。

$$\psi_X^{(n)}(0) = E[X^n] \quad (n \geq 1)$$

が成立する。モーメント母関数から分布が得られることが知られている。
確率変数 X の特性関数は

$$\phi(t)_X = E[e^{itX}]$$

で定義される。

example 確率密度関数が

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \frac{-(x-\mu)^2}{2\sigma^2}$$

で与えられる分布を平均 μ 分散 σ^2 の正規分布 $N(\mu, \sigma^2)$ という。X をこの確率変数とすると

$$\psi_X(t) = \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right)$$

X, Y をそれぞれ $N(\mu_1, \sigma_1^2), N(\mu_2, \sigma_2^2)$ の独立な確率変数とすると

$$\psi_{X+Y}(t) = E[e^{t(X+Y)}] = E[e^{tX}]E[e^{tY}] = \exp\left((\mu_1+\mu_2)t + \frac{1}{2}(\sigma_1^2+\sigma_2^2)t^2\right)$$

したがって、$X+Y$ は $N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$ の確率変数となる。

1.5 条件確率

X, Y を離散的確率変数とする。$Y=y$ のときの X の条件確率を

$$P\{X = x | Y = y\} = \frac{P\{X = x, Y = y\}}{P\{Y = y\}}$$

によって定義する。$Y = y$ のときの X の分布は

$$F(x|y) = P\{X \leq x | Y = y\}$$

で $Y=y$ のときの X の期待値は

$$E[X|Y=y]=\int_{\infty}^{\infty}xdF(x|y)=\sum_{x}xP\{X=x|Y=y\}$$

によって定義する。

X,Y の合成確率密度関数を $f(x,y)$ とするとき、条件確率密度は

$$f(x|y)=\frac{f(x,y)}{f_Y(y)}$$

で、この分布は

$$F(x|y)=\int_{-\infty}^{x}f(x|y)dx$$

定義される。$E[X|Y]$ を $Y=y$ のとき $E[X|Y=y]$ の値をとる確率変数とすると

$$E[X]=E[E[X|Y]]=\int_{-\infty}^{\infty}E[X|Y=y]dF_Y(y)$$

なぜなら、

$$\begin{aligned}\sum_{y}E[X|Y=y]P\{Y=y\}&=\sum_{y}\sum_{x}xP\{X=x|Y=y\}P\{Y=y\}\\&=\sum_{y}\sum_{x}xP\{X=x,Y=y\}\\&=\sum_{x}xP\{X=x\}\\&=E[X]\end{aligned}$$

A を event, $1_A(x)$ を A の indicator function とする。

$$1_A(x)=\begin{cases}1 & x\in A\\0 & x\notin A\end{cases}$$

このとき、

$$P\{A\}=E[1_A]=\int_{-\infty}^{\infty}P\{A|Y=y\}dF_Y(y)$$

X,Y を分布がそれぞれ F,G である独立な確率変数とすると、$X+Y$ の分布はつぎに定義される合成積 $F*G$ になる。

$$(F*G)(a)=P\{X+Y\leq a\}=\int_{-\infty}^{\infty}F(a-y)dG(y)$$

example $X_1,X_2,\ldots$ を独立な同じ分布の確率変数とする。N をそれらに独立な、非負整数値をとる確率変数とする。$N=n$ を条件とする確立変数 $Y=\sum_{i=1}^{N}X_i$ のモーメント母関数は

$$E\left[\exp\left(t\sum_{i=1}^{N}X_i\right)\middle|N=n\right]=E\left[\exp\left(t\sum_{i=1}^{n}X_i\right)\right]=(\psi_X(t))^n$$

したがって

$$E\left[\exp\left(t\sum_{i=1}^{N} X_i\right)\middle| N\right] = (\psi_X(t))^N$$

$$\psi_Y(t) = E\left[\exp\left(t\sum_{i=1}^{N} X_i\right)\right] = E[(\psi_X(t))^N]$$

これを微分して

$$\psi'_Y(t) = E[N(\psi_X(t))^{N-1}\psi'_X(t)]$$
$$\psi''_Y(t) = E[N(N-1)(\psi_X(t))^{N-2}(\psi'_X(t))^2 + N(\psi_X(t))^{N-1}\psi''_X(t)]$$

これより

$$E[Y] = E[N]E[X], \quad E[Y^2] = E[N]Var(X) + E[N^2]E^2[X]$$

であるから

$$Var(Y) = E[N]Var(X) + E^2[X]Var(N)$$

問題　つぎに与えられる分布を 2 項分布という。

$$\binom{n}{x} p^x (1-p)^{n-x} \qquad (x = 0, 1, \ldots, n)$$

ただし $0 \le p \le 1$ とする。モーメント母関数、平均、分散をもとめよ。

問題　つぎに与えられる分布を Poisson 分布という。

$$e^{-\lambda}\frac{\lambda^x}{x!} \qquad (x = 0, 1, \ldots, n)$$

ただし $0 < \lambda$ とする。モーメント母関数、平均、分散をもとめよ。

1.6　指数分布

つぎの確率密度 $f(x)$ をもつ確率変数 X は指数分布 $F(x)$ をもつという。

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$$F(x) = \int_{-\infty}^{x} f(y)dy = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

moment 母関数は

$$E[e^{tX}] = \int_0^\infty e^{\lambda x} \lambda e^{-\lambda x} dx = \frac{\lambda}{\lambda - x}$$

これより

$$E[X] = \frac{1}{\lambda}, \qquad Var(X) = \frac{1}{\lambda^2}$$

X が memoryless であるとは

$$P\{X > s + t | X > t\} = P\{X > s\} \qquad (s, t \geq 0)$$

が成立するときにいう。前の状態を忘れるということ。これを書き換えれば

$$\overline{F}(x+t) = \overline{F}(s)\overline{F}(t)$$

$\overline{F}(x)$ が memoryless であるとと、X が指数分布の確率変数であることとは同値である。($\overline{F}$ の連続性が必要。) 実際 $\overline{F}$ が上式をみたせば $\overline{F}(x) = e^{-\lambda x}$ となる定数 $\lambda > 0$ が存在するからである。

　分布関数 $F(x)$, 確率密度 $f(x)$ をもつ確率変数 X に対して

$$\lambda(t) = \frac{f(t)}{\overline{F}(t)}$$

を hazard rate 関数という。

$$P\{X \in (t, t+dt) | X > t\} \approx \lambda(t)dt$$

が成立する。

　指数分布においては hazard rate 関数は定数である。

　$\lambda(t)$ によって $\overline{F}(t)$ が決まる。実際

$$\lambda(t) = -\frac{d \log \overline{F}(t)}{dt}$$

と $\overline{F}(0) = 1$ より $\overline{F}(t)$ が得られる。

1.7 Limit theorem

大数の強法則　$X_1, X_2, \ldots$ を独立な同じ μ をもつ分布の確率変数とするとき

$$P\left\{\lim_{n\to\infty}\frac{X_1+X_2+\cdots+X_n}{n}=\mu\right\}=1$$

(たくさんの試行で平均をとれば平均値に近づく)
これはつぎのすこし一般なかたちで示すことができる。

$X_1, X_2, \ldots$ の和の分散を σ_n^2 とすれば、もし、

$$\frac{\sigma_n^2}{n^2}\to 0\ (n\to\infty)$$

ならば、任意の $\epsilon>0$ に対して

$$P\left\{\left|\frac{X_1+\cdots+X_n}{n}-\frac{E[X_1]+\cdots+E[X_n]}{n}\right|>\epsilon\right\}\to 0\ (n\to\infty)$$

が成立する。証明にはつぎのチェビシェフの不等式を用いる。

$$P\left\{\left|\frac{X_1+\cdots+X_n}{n}-\frac{E[X_1]+\cdots+E[X_n]}{n}\right|>\epsilon\right\}\le\frac{\sigma_n^2}{\epsilon^2n^2}$$

中心極限定理　$X_1, X_2, \ldots$ を独立な同じ平均 μ、同じ分散 σ^2 をもつ分布の確率変数とするとき

$$\lim_{n\to\infty}P\left\{\frac{X_1+X_2+\cdots+X_n-n\mu}{\sigma\sqrt{n}}\le a\right\}=\int_{-\infty}^{a}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx$$

(繰り返せば、和は正規分布に近づく)

すこし一般的につぎを証明しよう。$X_1, X_2, \ldots$ を独立確率変数で、$E[X_i]=0$, $V(X_i)$ は有限とする。また $\sigma_n^2=V(X_1+\cdots+X_n)$ とする。任意の $\epsilon>0$ に対し

$$\frac{1}{\sigma_n^2}\sum_{k=1}^{n}\int_{-\epsilon\sigma_n}^{\epsilon\sigma_n}x^2dF_k(x)\to 1\ (n\to\infty)$$

を Lindeberg の条件という。$X_1, X_2, \ldots$ が Lindeberg 条件をみたすならば、$\dfrac{X_1+X_2+\cdots+X_n}{\sigma_n}$ の分布関数は $N(0,1)$ に収束する。

証明　任意の $\epsilon>0$ に対して

$$\sum_{k=1}^{n}\int_{|x|>\epsilon\sigma_n}dF_k(x)\le\frac{1}{\epsilon^2\sigma_n}\sum_{k=1}^{n}\int_{|x|>\epsilon\sigma_n}x^2dF_k(x)$$

$$= \frac{1}{\epsilon^2 \sigma_n} \left\{ 1 - \frac{1}{V_n} \sum_{k=1}^{n} \int_{|x| \le \epsilon \sigma_n} x^2 dF_k(x) \right\}$$

$$\to 0 \quad (n \to \infty)$$

$\exp \dfrac{itx}{\sigma_n}$ の展開から

$$\exp \frac{itx}{\sigma_n} = 1 + \frac{itx}{\sigma_n} - \frac{\theta t^2 x^2}{2\sigma_n} \qquad (|\theta| \le 1)$$

$F_k(x)$ の特性関数を $f_k(t)$ とおけば、$E[X_k] = 0$ であるから

$$\sum_{k=1}^{n} \left| 1 - f_k\left(\frac{t}{\sigma}\right) \right| = \sum_{k=1}^{n} \left| \int_{-\infty}^{\infty} \left(1 - \exp \frac{itx}{\sigma_n} \right) dF_k(x) \right| \le \frac{t^2}{2}$$

となる。また、$\left| 1 - \exp \dfrac{itx}{\sigma_n} \right| \le \dfrac{|tx|}{\sigma_n}$ であるから

$$\left| 1 - f_k\left(\frac{t}{\sigma}\right) \right| \le \left| \int_{|x| > \epsilon\sigma_n} \right| + \left| \int_{|x| \le \epsilon\sigma_n} \right| \le 2 \int_{|x| > \epsilon\sigma_n} dF_k(x) + |t|\epsilon$$

よって $T > 0$ に対して $|t| \le T$ で一様に

$$\max_{1 \le i \le n} \left| 1 - f_k\left(\frac{t}{\sigma_n}\right) \right| \to 0 \qquad (n \to \infty)$$

また

$$\left| \log \prod_{k=1}^{n} f_k\left(\frac{t}{\sigma_n}\right) - \sum_{k=1}^{n} \left(f_k\left(\frac{t}{\sigma_n}\right) - 1 \right) \right|$$

$$\le \frac{1}{2} \frac{\max_k \left| f_k\left(\frac{t}{\sigma_n}\right) - 1 \right|}{1 - \max_k \left| f_k\left(\frac{t}{\sigma_n}\right) - 1 \right|} \sum_{k=1}^{n} \left| f_k\left(\frac{t}{\sigma_n}\right) - 1 \right|$$

($\log x$ の展開から) を得るから、$|t| \le T$ で一様に

$$\left| \log \prod_{k=1}^{n} f_k\left(\frac{t}{\sigma_n}\right) - \sum_{k=1}^{n} \left(f_k\left(\frac{t}{\sigma_n}\right) - 1 \right) \right| \to 0 \qquad (n \to \infty)$$

つぎに

$$f_k\left(\frac{tx}{\sigma_n}\right) - 1 = \int_{-\infty}^{\infty} \left(\exp \frac{itx}{\sigma_n} - 1 \right) dF_k(x) = \int_{|x| \le \epsilon\sigma_n} + \int_{|x| > \epsilon\sigma_n}$$

$$= \int_{|x| \le \epsilon\sigma_n} \left(\frac{itx}{\sigma_n} - \frac{t^2}{2} \frac{x^2}{\sigma_n^2} + \frac{it^3}{6} \frac{x^3}{\sigma_n^3} \theta \right) dF_k(x)$$

$$+ \int_{|x| > \epsilon\sigma_n} \left[\left(\frac{itx}{\sigma_n} - \frac{t^2}{2} \frac{x^2}{\sigma_n^2} \right) + \left(\frac{t^2}{2} \frac{x^2}{\sigma_n^2} + \frac{t^2}{2} \frac{x^2}{\sigma_n^2} \theta' \right) \right] dF_k(x)$$

$$= -\frac{t^2}{2}\frac{E[X_k^2]}{\sigma_n^2} + R_k \qquad (|\theta|, |\theta'| \le 1)$$

となる。ここに

$$R_k = \frac{t^3}{6}\int_{-\epsilon\sigma_n}^{\epsilon\sigma_n} \frac{x^3}{\sigma_n^3}\theta dF_k(x) + \frac{t^2}{2}\int_{|x|>\epsilon\sigma_n} \left(\frac{x^2}{\sigma_n^2} + \frac{x^2}{\sigma_n^2}\theta'\right) dF_k(x)$$

さて

$$|R_k| \le \frac{t^3}{6}\int_{-\epsilon\sigma_n}^{\epsilon\sigma_n} \epsilon\frac{E[X_k^2]}{\sigma_n^2} + \frac{t^2}{2}\frac{2}{\sigma_n^2}\int_{|x|>\epsilon\sigma_n} x^2 dF_k(x)$$

であるから

$$\sum_{k=1}^{n}\left(f_k\left(\frac{t}{\sigma_n}\right) - 1\right) + \frac{t^2}{2} = \sum_{k=1}^{n} R_k$$

は $|t| \le T$ において一様に 0 に近づく。よって、

$$\prod_{k=1}^{n} f_k\left(\frac{t}{\sigma_n}\right) \to \exp\left(-\frac{t^2}{2}\right) \qquad (t \to \infty)$$

右辺は $N(0,1)$ の特性関数であるから求める結果を得る。

問題　Polya の壷。b この黒球と r この赤球がはいっている壷がある。1 こ任意にとりだし、取り出された色と同色の c こをつけくわえて戻す。これを繰り返す。最初より n 番目までの取り出しで k この黒球が出る確率を $p_k(n)$ で表すと、

$$p_k(n+1) = p_k(n)\frac{r+(n-k)c}{b+r+nc} + p_{k-1}(n)\frac{b+(k-1)c}{b+r+nc}$$

が成立する。ただし $p_{-1}(n) = 0$ とする。

2 Poisson Process

2.1 Poisson Process

集合 T の各元 t に対し $X(t)$ を確率変数とするとき、$\underline{X} = \{X(t) \mid t \in T\}$ を確率過程という。T が可算であるとき、それは離散過程、連続体であるとき、連続過程という。

連続過程（微分可能性を仮定する）$\{X(t) \mid t \in T\}$ は、任意の時間列
$t_0 < t_1 < \cdots < t_n$ に対して

$$X(t_1) - X(t_0), X(t_2) - X(t_1), \ldots, X(t_n) - X(t_{n-1})$$

が独立であるとき、独立な増分をもつという。$X(t+s) - X(t)$ が任意の t に対して同じ分布を与えるとき、この確率過程は定常増分をもつという。

確率過程 $N(t) \mid t \geq 0\}$ が計数過程であるとはつぎをみたすときにいう。

(i) $N(t) \geq 0$

(ii) $N(t)$ は整数値をとる

(iii) もし $s < t$ ならば $N(s) \leq N(t)$

(iv) $s < t$ ならば $N(t) - N(s)$ は区間 $(s, t]$ に起きた事象の個数を示す

計数過程は、互いに交わらない区間に起きた事象の個数が独立であるとき、独立な増分をもつという。また、それが定常増分をもつとは、任意の時間区間に起きた事象の個数がその区間のみに依存するときにいう。すなわち、任意の $t_1 < t_2, s > 0$ に対して $N(t_2 + s) - N(t_1 + s)$ と $N(t_2) - N(t_1)$ とが同じ分布をもつということである。

定義　つぎの計数過程 $N(t)$ は比率 $\lambda > 0$ をもつ Poisson 過程といわれる。

(i) $N(0) = 0$

(ii) $N(t)$ は独立な増分をもつ

(iii) もし $s, t > 0$ ならば
$$P\{N(t+s) - N(s) = n\} = e^{-\lambda t}\frac{(\lambda t)^n}{n!}, \quad n = 0, 1, \ldots$$

(iii) から Poisson 過程は定常増分をもち、

$$E[N(t)] = \lambda t$$

別定義　つぎの計数過程 $N(t)$ は比率 $\lambda > 0$ をもつ Poisson 過程といわれる。

(i)　$N(0) = 0$

(ii)　$N(t)$ は定常的で独立な増分をもつ

(iii)　$P\{N(h) = 1\} = \lambda h + o(h)$

(iv)　$P\{N(h) \geq 2\} = o(h)$

定理　上の２つの定義は同値である。

Proof　別定義が定義を導くことのみ示す。逆は演習とする。

$$P_n(t) = P\{N(t) = n\}$$

とおく。

$$\begin{aligned} P_0(t+h) &= P\{N(t+h) = 0\} \\ &= P\{N(t) = 0, N(t+h) - N(t) = 0\} \\ &= P\{N(t) = 0\} P\{N(t+h) - N(t) = 0\} \\ &= P_0(t)(1 - \lambda h + o(h)) \end{aligned}$$

より

$$\frac{P_0(t+h) - P_0(t)}{h} = -\lambda P_0(t) + \frac{o(h)}{h}$$

そして

$$P_0'(t) = -\lambda P_0(t)$$

これより、$P_0(0) = 1$ に考慮して

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

つぎに $n \geq 1$ として

$$\begin{aligned} P_n(t+h) = & \ P\{N(t+h) = n\} \\ = & \ P\{N(t) = n, N(t+h) - N(t) = 0\} \\ + & \ P\{N(t) = n-1, N(t+h) - N(t) = 1\} \\ + & \ P\{N(t+h) = n, N(t+h) - N(t) \geq 2\} \end{aligned}$$

よって

$$\begin{aligned} P_n(t+h) = & \ P_n(t) + P_0(h) + P_{n-1}(t) P_1(h) + o(h) \\ = & \ (1 - \lambda h) P_n(t) + \lambda h P_{n-1}(T) + o(h) \end{aligned}$$

これより、前と同様にして

$$P_n'(t) = \lambda P_n(t) + \lambda P_{n-1}(t)$$

すなわち

$$\frac{d}{dt}(e^{\lambda t}P_n(t)) = \lambda e^{\lambda t}P_{n-1}(t)$$

帰納法によって

$$P_n(t) = e^{-\lambda t}\frac{(\lambda t)^n}{n!}$$

を得る。実際

$$\frac{d}{dt}(e^{\lambda t}P_n(t)) = \frac{\lambda(\lambda t)^n}{(n-1)!}$$

とすれば、

$$e^{\lambda t}P_n(t) = \frac{(\lambda t)^n}{n!} + \mathrm{const.}$$

を得、$P_n(0) = P\{N(0) = n\} = 0$ より、$\mathrm{const.} = 0$ を得る。

2.2 interarrival and waiting times

$\{N(t) \mid t \geq 0\}$ を Poisson 過程とする。X_1 を第1回目 event が起きる時間、X_n を第 $n-1$ 回目 event のつぎに第 n 回目 event が起きる時間とする。$\{X_n \mid n \geq 1\}$ は interarrival times 列とよばれる。

それらの分布はつぎのようになる。

$$P\{X_1 > t\} = e^{-\lambda t}$$

$0 \leq s_1 < s_2 < \cdots < s_{n-1} < t$ とするとき

$$P\{X_n > t \mid X_1 = s_1, X_2 = s_2, \ldots, X_{n-1} = s_{n-1}\} = e^{-\lambda t}$$

すなわち、$\{X_n\}$ は独立な平均値 $1/\lambda$ をもつ同型の指数的確率変数である。実際、

$$P\{X_1 > t\} = P\{N(t) = 0\} = e^{-\lambda t}$$

である。独立増分と定常増分によって

$$P\{X_2 > t \mid X_1 = s\} = P\{N(s+t) = 1 \mid N(s) = 1\} = P\{N(s+t) = N(s)\} = P\{N(t) = 0\} = e^{-\lambda t}$$

以下同様。

$S_n = \sum_{i=1}^{n} X_i$ は n 回目 event がおきるまでの待ち時間である。S_n の分布はガンマ分布といわれる。

$$f(t) = \frac{d}{dt} P\{S_n \leq t\} = \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} \qquad (t > 0)$$

実際

$$N(t) \geq n \Longleftrightarrow S_n \leq t$$

であるから

$$P\{S_n \leq t\} = P\{N(t) \geq n\} = \sum_{j=n}^{\infty} e^{-\lambda t} \frac{(\lambda t)^j}{j!}$$

を微分すれば、

$$f(t) = -\sum_{j=n}^{\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^j}{j!} + \sum_{j=n}^{\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^{j-1}}{(j-1)!} = \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} \qquad (t > 0)$$

2.3 arrival times の条件分布

$s \le t$ とするとき

$$P\{X_1 < s \mid N(t) = 1\} = \frac{P\{X_1 < s, N(t) = 1\}}{P\{N(t) = 1\}} = \frac{P\{[0,s) \text{ で } 1 \text{ 回 }, [s,t) \text{ で } 0 \text{ 回 }\}}{P\{N(t) = 1\}}$$

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

これは一様分布を示す。この結果を一般化する。

$Y_1, \ldots, Y_n$ を n この確率変数とする。$Y_{(k)}$ は $Y_1, \ldots, Y_n$ のなかで k 番目の最小値を表すとき、$Y_{(1)}, \ldots, Y_{(n)}$ を $Y_1, \ldots, Y_n$ の順序統計という。もし、Y_i らが独立で同型の連続密度関数 f をもつならば、$Y_{(1)}, \ldots, Y_{(n)}$ の結合密度関数はつぎによって与えられる。

$$f(y_1, \ldots, y_n) = n! \prod_{i=1}^{n} f(y_i)$$

y_i がすべて異なる場合に、これは成立する。あとは連続性から OK。もし、Y_i らが $(0,t)$ 上で一様ならば

$$f(y_1, \ldots, y_n) = \frac{n!}{t^n}$$

定理　$N(t) = n$ とせよ。このとき $S_1, \ldots, S_n$ は $(0,t)$ 上 n この独立な一様分布の順序統計に一致する。

証明　$0 < t_1 < \cdots < t_n < t_{n+1} = t$ とする。$0 < h_i$ を十分小さくとって $t_i + h_i < t_{i+1}$ が成立するとする。すると

$$P\{t_i \le S_i \le t_i + h_i \ (1 \le \le n) \mid N(t) = n\}$$

$$= \frac{P\{[t_i, t_i + h_i] \text{ でちょうど } 1 \text{ 回起きる } \ (1 \le i \le n), [0,t] \text{ で起きない }\}}{P\{N(t) = n\}}$$

$$= \frac{\lambda h_1 e^{-\lambda h_1} \cdots \lambda h_n e^{-\lambda h_n} e^{-\lambda(t-h_1-h_2-\cdots-h_n}}{e^{-\lambda t}(\lambda t)^n / n!} = \frac{n!}{t^n} h_1 h_2 \cdots h_n$$

つぎのような問題を考えよう。率 λ のポアソン過程に基づいて客が到達する。列車が時刻 t で出発するとき時間 t までに到達する客の待ち時間の平均はどれほどか？すなわち $E\left[\sum_{i=1}^{N(t)} (t - S_i)\right]$ を計算することになる。

$$E\left[\sum_{i=1}^{N(t)} (t - S_i) \mid N(t) = n\right] = E\left[\sum_{i=1}^{n} (t - S_i) \mid N(t) = n\right] = nt - E\left[\sum_{i=1}^{n} S_i \mid N(t) = n\right]$$

いま $U_1, \dots, U_n$ を n この独立な $(0,t)$ における一様確率変数とすると、

$$E\left[\sum_{i=1}^{n} S_i \mid N(t)=n\right] = E\left[\sum_{i=1}^{n} U_{(i)}\right] = E\left[\sum_{i=1}^{n} U_i\right] = \frac{nt}{2}$$

したがって

$$E\left[\sum_{i=1}^{N(t)} (t-S_i) \mid N(t)=n\right] = \frac{t}{2}E[N(t)] = \frac{\lambda t^2}{2}$$