

## 1.8 確率変数の不等式 1

定理 1.9 (マルコフの不等式) 非負値確率変数  $X$  が有限の期待値  $\mathbb{E}[X] < \infty$  をもつとき, 任意の正数  $a > 0$  に対して,

$$P\{X \geq a\} \leq \frac{1}{a}\mathbb{E}[X]$$

が成立する.

証明 証明  $X$  が連続型確率の場合を示す.  $X$  の確率密度関数を  $f_X(x)$  とする.

$$\mathbb{E}[X] = \int_0^{\infty} x f_X(x) dx \geq \int_a^{\infty} x f_X(x) dx \geq a \int_a^{\infty} f_X(x) dx = a P(X \geq a)$$

よりわかる.

証明

定理 1.10 (チェビシエフの不等式) 確率変数  $X$  が平均  $\mu$ , 分散  $\sigma^2$  をもつとき, 任意の正数  $a > 0$  に対して,

$$\mathbb{P}\{|X - \mu| \geq a\} \leq \frac{\sigma^2}{a^2}$$

が成り立つ.

証明 Markov の不等式から

$$\mathbb{P}\{|X - \mu| \geq a\} = \mathbb{P}\{|X - \mu|^2 \geq a^2\} \geq \frac{\mathbb{E}[(X - \mu)^2]}{a^2}$$

よりわかる. □

例 1.13  $X$  は平均  $\mu$ , 分散  $\sigma^2$  ( $\sigma > 0$ ) の分布に従うとする. このとき, 任意の正の実数  $t$  に対して,

$$\mathbb{P}\{|X - \mu| \geq t\sigma\} \leq \frac{1}{t^2} \mathbb{E}\left[\left(\frac{|X - \mu|}{\sigma}\right)^2\right] = \frac{1}{t^2} \frac{\mathbb{E}[(X - \mu)^2]}{\sigma^2} = \frac{1}{t^2}$$

となる. したがって,  $t = 2$  の場合

$$\mathbb{P}\{|X - \mu| \geq 2\sigma\} \leq \frac{1}{2^2} = 0.25$$

となる.

例 1.14  $Z$  は標準正規分布に従うとする. このとき, 任意の正の実数  $t$  に対して,

$$\mathbb{P}\{|Z| \geq t\} \leq \sqrt{\frac{2}{\pi}} \frac{e^{-t^2/2}}{t} \tag{1.2}$$

となる.

$t = 2$  のとき, Chebychev の不等式から

$$\mathbb{P}\{|Z| \geq 2\} \leq \frac{1}{2^2} = 0.25$$

となる. しかし, (1.2) から

$$\mathbb{P}\{|Z| \geq 2\} \leq \sqrt{\frac{2}{\pi}} \frac{e^{-2}}{2} = 0.054$$

となる. また,

$$\mathbb{P}\{|Z| \geq 3\} \leq \sqrt{\frac{2}{\pi}} \frac{e^{-4.5}}{3} = 0.00295$$

となる。

1.2) は以下からわかる。

$$\begin{aligned}\mathbb{P}\{Z \geq t\} &= \frac{1}{\sqrt{2\pi}} \int_t^\infty e^{-z^2/2} dz \\ &\leq \frac{1}{\sqrt{2\pi}} \int_t^\infty \frac{z}{t} e^{-z^2/2} dz \\ &= \frac{1}{\sqrt{2\pi}} \left[ -\frac{1}{t} e^{-z^2/2} \right]_t^\infty = \frac{1}{\sqrt{2\pi}} \frac{e^{-t^2/2}}{t}\end{aligned}$$

と

$$\mathbb{P}\{|Z| \geq t\} = \mathbb{P}\{\{Z \geq t\} \cup \{Z \leq -t\}\} = \mathbb{P}\{Z \geq t\} + \mathbb{P}\{Z \leq -t\} = 2\mathbb{P}\{Z \geq t\}$$

からわかる。

ある区間  $I = (a, b)$  の上の実数値連続関数  $h(x)$  が凸 (convex) であるとは、任意の  $c \in (0, 1)$  と任意の  $x_1, x_2 \in I$  に対して、

$$h(cx_1 + (1-c)x_2) \leq ch(x_1) + (1-c)h(x_2)$$

が成り立つことである。もし、 $x_1 \neq x_2$  に対して「 $\leq$ 」のかわりに「 $<$ 」が常に成立するならば、狭義の凸関数という。

**定理 1.11 (イェンセンの不等式)**  $X$  を確率変数とし、 $h(x)$  を  $X$  の値域を含む区間上で凸な関数とする。 $X$  と  $h(X)$  の期待値が有限のとき、

$$\mathbb{E}[h(X)] \geq h(\mathbb{E}[X])$$

が成立する。もし、 $h(x)$  が狭義の凸関数のとき、等号が成り立つのは1点分布の時に限られる。

**証明** 任意の固定した  $x_0$  とある線形関数  $g(x) = ax + b$  が存在して、 $h(x_0) = g(x_0)$  とすべての  $x$  に対して  $h(x) \geq g(x)$  が成立する。 $x_0 = \mathbb{E}[X]$  として、上のことを利用すれば、

$$\begin{aligned}\mathbb{E}[h(X)] &\geq \mathbb{E}[g(X)] \\ &= \mathbb{E}[aX + b] \\ &= a\mathbb{E}[X] + b \\ &= g(\mathbb{E}[X]) = h(\mathbb{E}[X])\end{aligned}$$

□

**例 1.15**  $X$  を  $\mathbb{E}[X^2] < \infty$  なる確率変数とする。Jensen の不等式を利用するために、まず、 $\mathbb{E}[|X|] < \infty$  を確認する。たとえば、 $X$  が確率密度関数  $f_X(x)$  を持てば

$$\begin{aligned}\mathbb{E}[|X|] &= \int_{-\infty}^{\infty} |x| f_X(x) dx \\ &= \int_{-\infty}^{-1} |x| f_X(x) dx + \int_{-1}^1 |x| f_X(x) dx + \int_1^{\infty} |x| f_X(x) dx \\ &\leq \int_{-\infty}^{-1} x^2 f_X(x) dx + \int_{-1}^1 f_X(x) dx + \int_1^{\infty} x^2 f_X(x) dx \\ &\leq \int_{-\infty}^{\infty} |x|^2 f_X(x) dx + \int_{-\infty}^{\infty} f_X(x) dx = \mathbb{E}[X^2] + 1 < \infty\end{aligned}$$

となる． $X$  が離散型確率変数のときも同様にすれば， $\mathbb{E}[|X|] < \mathbb{E}[X^2] + 1$  であることがわかる．さらに， $\mathbb{E}[|X|]$  の評価が Jensen の不等式から得られる． $h(x) = x^2$  とすれば， $h(x)$  は狭義の凸関数なので，Jensen の不等式より

$$\{\mathbb{E}[X]\}^2 = g(\mathbb{E}[X]) \leq \mathbb{E}[g(X)] = \mathbb{E}[X^2]$$

となる．したがって，

$$|\mathbb{E}[X]| \leq \sqrt{\mathbb{E}[X^2]}$$

をえる．

例 1.16  $a_1, a_2, \dots, a_n$  は正の数とし，

$$\begin{aligned} m_A &= \frac{1}{n}(a_1 + a_2 + \dots + a_n) && \text{(算術平均)} \\ m_G &= \{a_1 a_2 \dots a_n\}^{1/n} && \text{(幾何平均)} \\ m_H &= \frac{1}{\frac{1}{n} \left( \frac{1}{a_1} + \frac{1}{a_2} + \dots + \frac{1}{a_n} \right)} && \text{(調和平均)} \end{aligned}$$

とする．このとき，

$$m_H \leq m_G \leq m_A$$

が成立する．

Jensen の不等式を用いて示す：そのために， $X$  を確率変数として

$$\mathbb{P}\{X = a_i\} = \frac{1}{n}, \quad i = 1, 2, \dots, n$$

とする． $-\log x$  は凸関数なので， $\mathbb{E}[\log X] \leq \log(\mathbb{E}[X])$  が成立する．したがって，

$$\log m_G = \frac{1}{n} \sum_{i=1}^n \log a_i = \mathbb{E}[\log X] \leq \log(\mathbb{E}[X]) = \log \left( \frac{1}{n} \sum_{i=1}^n a_i \right) = \log m_A$$

より  $m_G \leq m_A$  がわかる．また，

$$\log \frac{1}{m_H} = \log \left( \frac{1}{n} \sum_{i=1}^n \frac{1}{a_i} \right) = \log \left( \mathbb{E} \left[ \frac{1}{X} \right] \right) \geq \mathbb{E} \left[ \log \left( \frac{1}{X} \right) \right] = -\mathbb{E}[\log X] = -\log m_G$$

から  $m_G \geq m_H$  がわかる．

