

チェビシェフの不等式 と 大数の法則 (高校生に向けて)

確率分布の平均（期待値）、分散などの基本性質から、チェビシェフの不等式、大数の法則までを高校生に向けて、ていねいに説明・証明をした。具体的には、計算過程を急がず、一段一段確認しながら進める。

[記号] 一般的な記号は既知とするが、確認のために次を記す。

事象 A の起こる確率 $P(A)$, 確率変数 X

平均（期待値） $m = E(X)$, 分散 $\sigma^2 = V(X)$, 標準偏差 σ

[準備 1] 確率分布が右の表の場合

$$P(x_i) = p_i \quad (i=1, 2, \dots, n)$$

X	x_1	x_2	\dots	x_n	計
P	p_1	p_2	\dots	p_n	1

$$\sum_{i=1}^n p_i = 1, \quad m = E(X) = \sum_{i=1}^n p_i x_i$$

$$V(X) = E\left((X - m)^2\right) = \sum_{i=1}^n p_i (x_i - m)^2 = E(X^2) - \{E(X)\}^2 \dots \textcircled{1}, \quad \sigma = \sqrt{V(X)}$$

[証明 ①]

$$\begin{aligned} V(X) &= \sum_{i=1}^n p_i (x_i - m)^2 = \sum_{i=1}^n (p_i x_i^2 - 2p_i m x_i + p_i m^2) = \sum_{i=1}^n p_i x_i^2 - 2m \sum_{i=1}^n p_i x_i + m^2 \sum_{i=1}^n p_i \\ &= E(X^2) - 2m^2 + m^2 = E(X^2) - m^2 = E(X^2) - \{E(X)\}^2 \quad \blacksquare \end{aligned}$$

特に確率分布が右の表の場合

$$E(X) = pa + qb = pa + (1-p)b$$

X	a	b	計
P	p	$q (=1-p)$	1

この簡単な等式は、後に重要な箇所でも参照することになる。

[準備 2] a, b を定数とするとき、

(1) 確率変数 X, Y に対して、平均 $E(X)$ の線形性より

$$E(X + Y) = E(X) + E(Y), \quad E(aX + b) = aE(X) + b$$

となる。さらに $V(aX + b) = a^2 V(X) \dots \textcircled{2}$

(2) 互いに独立な確率変数 X, Y に対して

$$E(XY) = E(X)E(Y) \dots \textcircled{3}$$

$$V(X + Y) = V(X) + V(Y) \dots \textcircled{4}$$

[証明②, ③, ④]

$$\begin{aligned} V(aX+b) &= E\left((aX+b)^2\right) - \{E(aX+b)\}^2 = E(a^2X^2 + 2abX + b^2) - \{aE(X)+b\}^2 \\ &= a^2E(X^2) + 2abE(X) + b^2 - a^2\{E(X)\}^2 - 2abE(X) - b^2 \\ &= a^2\left(E(X^2) - \{E(X)\}^2\right) = a^2V(X) \end{aligned}$$

$X = \{x_1, \dots, x_n\}$, $Y = \{y_1, \dots, y_m\}$, $P(x_i) = p_i (i=1, \dots, n)$, $P(y_j) = q_j (j=1, \dots, m)$ とする。

$$\begin{aligned} E(XY) &= \sum_{j=1}^m \sum_{i=1}^n (p_i q_j \cdot x_i y_j) = \sum_{j=1}^m \sum_{i=1}^n (q_j y_j \cdot p_i x_i) = \sum_{j=1}^m \left\{ q_j y_j \sum_{i=1}^n (p_i x_i) \right\} = \sum_{j=1}^m \{q_j y_j E(X)\} \\ &= E(X) \sum_{j=1}^m (q_j y_j) = E(X)E(Y) \end{aligned}$$

$$\begin{aligned} V(X+Y) &= E\left((X+Y)^2\right) - \{E(X+Y)\}^2 = E(X^2 + 2XY + Y^2) - \{E(X) + E(Y)\}^2 \\ &= \left(E(X^2) + 2E(XY) + E(Y^2)\right) - \left(\{E(X)\}^2 + 2E(X)E(Y) + \{E(Y)\}^2\right) \\ &= \left(E(X^2) - \{E(X)\}^2\right) + \left(E(Y^2) - \{E(Y)\}^2\right) = V(X) + V(Y) \quad \blacksquare \end{aligned}$$

[準備3] 確率変数とその平均の定義から明らか。説明の必要は無いだろう。

すべての x に対して $f(x) \geq g(x)$ ならば $E(f(X)) \geq E(g(X))$ である。

特に、等号が成り立つ条件は、 $P(f(X) = g(X)) = 1$ である。

[定理1]

X_1, X_2, \dots, X_n 互いに独立な確率変数で、すべて同じ分布に従っているとす。つまり、

$$E(X_1) = E(X_2) = \dots = E(X_n) = m (< \infty)$$

$$V(X_1) = V(X_2) = \dots = V(X_n) = \sigma^2 (< \infty)$$

である。そして、

$$\bar{X} = \frac{1}{n}(X_1 + X_2 + \dots + X_n)$$

とする。 \bar{X} を n 個の確率変数の平均であることを明示するため $\bar{X} = \bar{X}_{(n)}$ とかく。

このとき、 $E(\bar{X}_{(n)}) = m \dots$ ⑤, $V(\bar{X}_{(n)}) = \frac{\sigma^2}{n} \dots$ ⑥ である。

[証明⑤, ⑥]

$$E(\bar{X}_{(n)}) = E\left(\frac{1}{n}(X_1 + X_2 + \dots + X_n)\right) = \frac{1}{n}E(X_1 + X_2 + \dots + X_n)$$

$$= \frac{1}{n}\{E(X_1) + E(X_2) + \dots + E(X_n)\} = \frac{1}{n}(m + m + \dots + m) = \frac{nm}{n} = m$$

$$V(\bar{X}_{(n)}) = V\left(\frac{1}{n}(X_1 + X_2 + \dots + X_n)\right) = \frac{1}{n^2}V(X_1 + X_2 + \dots + X_n)$$

$$= \frac{1}{n^2}\{V(X_1) + V(X_2) + \dots + V(X_n)\} = \frac{1}{n^2}(\sigma^2 + \sigma^2 + \dots + \sigma^2) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n} \quad \blacksquare$$

[定理2] チェビシエフの不等式

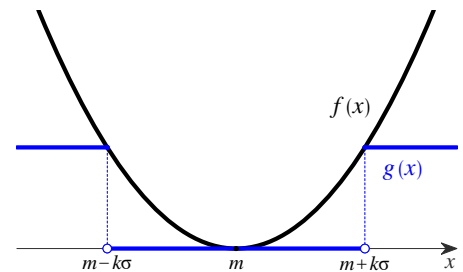
確率変数 X の平均 m , 標準偏差 σ に対して, 任意の $k(>0)$ に対して,

$$P(|X - m| \geq k\sigma) \leq \frac{1}{k^2} \dots \textcircled{7}$$

が成り立つ。

[証明⑦]

$$f(x) = \frac{1}{k^2}(x - m)^2, \quad g(x) = \begin{cases} \sigma^2 & (|x - m| \geq k\sigma) \\ 0 & (|x - m| < k\sigma) \end{cases}$$



とする。すべての x に対して $f(x) \geq g(x)$ であり, 準備3により次の不等式が得られる。

$$E(f(X)) \geq E(g(X)) \dots \textcircled{8}$$

ところで⑧の左辺は次のようになる。

$$E(f(X)) = E\left(\frac{1}{k^2}(X - m)^2\right) = \frac{1}{k^2}E((X - m)^2) = \frac{1}{k^2}V(X) = \frac{\sigma^2}{k^2} \dots \textcircled{9}$$

また, $P(|X - m| \geq k\sigma) = p$ とおけば $P(|X - m| < k\sigma) = 1 - p$ であるので, $g(x)$ の定義よりその分布は次の様になっている。

X	σ^2	0	計
P	$p (= P(X - m \geq k\sigma))$	$1 - p (= P(X - m < k\sigma))$	1

この確率分布の平均を求めると, ⑧の右辺は次のようである。

$$E(g(X)) = p \cdot \sigma^2 + (1 - p) \cdot 0 = p\sigma^2 = \sigma^2 P(|X - m| \geq k\sigma) \dots \textcircled{10}$$

⑧に⑨と⑩を代入し,

$$\frac{\sigma^2}{k^2} \geq \sigma^2 P(|X - m| \geq k\sigma) \quad \text{よって} \quad P(|X - m| \geq k\sigma) \leq \frac{1}{k^2} \quad \blacksquare$$

[定理 3] 大数の法則

任意の $\varepsilon (> 0)$ に対して, $n \rightarrow \infty$ のとき $P(|\bar{X} - m| \geq \varepsilon) \rightarrow 0 \cdots \textcircled{11}$

[証明 ⑪]

任意の ε に対し $k = \frac{\varepsilon}{\sigma}$ とおく。この k に対してチェビシエフの不等式 ⑦ は成り立つ。

さて, 上記と同様に $\frac{1}{n}(X_1 + X_2 + \cdots + X_n) = \bar{X} = \bar{X}_{(n)}$ とすると,

$$V(\bar{X}_{(n)}) = \frac{\sigma^2}{n}$$

であるので ⑨ と ⑩ は,

$$E(f(\bar{X}_{(n)})) = \frac{1}{k^2} V(\bar{X}_{(n)}) = \frac{\sigma^2}{nk^2}$$

$$E(g(\bar{X}_{(n)})) = \sigma^2 P(|\bar{X}_{(n)} - m| \geq k\sigma)$$

であるから,

$$\frac{\sigma^2}{nk^2} = E(f(\bar{X}_{(n)})) \geq E(g(\bar{X}_{(n)})) = \sigma^2 P(|\bar{X}_{(n)} - m| \geq k\sigma)$$

両辺を σ^2 で割り,

$$P(|\bar{X}_{(n)} - m| \geq k\sigma) \leq \frac{1}{nk^2} \cdots \textcircled{12}$$

を得る。ところで $k\sigma = \varepsilon$, $\frac{1}{k^2} = \frac{\sigma^2}{\varepsilon^2}$ であるので ⑫ に代入すると,

$$P(|\bar{X}_{(n)} - m| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2} = \frac{\sigma^2}{\varepsilon^2} \cdot \frac{1}{n}$$

$n \rightarrow \infty$ とすると $\frac{\sigma^2}{\varepsilon^2} \cdot \frac{1}{n} \rightarrow 0$ となるので,

$$n \rightarrow \infty \text{ のとき } P(|\bar{X} - m| \geq \varepsilon) = P(|\bar{X}_{(n)} - m| \geq \varepsilon) \rightarrow 0 \quad \blacksquare$$

※ (注意) 全体を通して, 平均や分散が有限な値で存在している事を前提としている。

[参考文献] 竹内 啓著「数理統計学」東洋経済新報社 昭和 38 年 (1963)

堀部和経 2022/06/10