

数理統計学まとめ (その4): 第3章 確率変数と確率分布

3 離散型確率変数の分布

いくつかの離散型確率変数の分布の例をあげよう.

1. (0-1) 分布: X は 0 か 1 の値を取る. その分布は

$$P(X = 1) = p \quad P(X = 0) = q (= 1 - p)$$

となる. 平均は p , 分散は $pq = p(1 - p)$ になる.

2. 2 項分布 $B(n, p)$: 同じ試行を (独立に) n 回繰り返す.

- 毎回の結果は事象 A が「起こる」(その確率 p) か「起こらない」(その確率 $q = 1 - p$) のどちらか.
- 毎回の結果は他の回の結果と独立.

このとき X を最終的に A が起こった回数とする. X の分布は

$$P(X = x) = \binom{n}{x} p^x q^{n-x} = \frac{n!}{x!(n-x)!} p^x q^{n-x}$$

となっている. $q = 1 - p$ である.

(n 回の試行のうち, どの x 回で起こるかその場合の数は $\binom{n}{x}$ 個. そのそれぞれの場合で A が指定された場所で x 回起こり, $n - x$ 回起こらない確率は $p^x q^{n-x}$.)

2 項分布の平均と分散: X を 2 項分布 $B(n, p)$ をもつ確率変数として,

$$E(X) = np, \quad V(X) = npq = np(1 - p)$$

が成り立つ. 証明は 2 項定理 による.

$$(a + b)^n = \sum_{x=0}^n \binom{n}{x} a^x b^{n-x}. \quad (1)$$

実際, X のとり得る値は 0 から n までの整数で,

$$\begin{aligned} E(X) &= \sum_{x=0}^n x \cdot \binom{n}{x} p^x q^{n-x} \\ &= \sum_{x=1}^n \frac{n!}{(x-1)!(n-x)!} p^x q^{n-x} \\ &= \sum_{z=0}^{n-1} np \frac{(n-1)!}{z!(n-1-z)!} \quad (z = x-1 \text{ と変数変換}) \\ &= np(p+q)^{n-1} = np \end{aligned}$$

$$E(X^2) = E(X(X-1) + X) = E(X(X-1)) + E(X)$$

$$\begin{aligned} &= \sum_{x=0}^n x(x-1) \binom{n}{x} p^x q^{n-x} + np \\ &= n(n-1)p^2 \sum_{x=2}^n \binom{n-2}{x-2} p^{x-2} q^{n-x} + np \\ &= n(n-1)p^2 + np \end{aligned}$$

$$V(X) = E(X^2) - (E(X))^2 = n(n-1)p^2 + np - n^2p^2 = np(1-p) = npq$$

3. ポアソン分布 $Po(\mu)$ μ というパラメータを一つ持つ. とり得る値は $0, 1, 2, \dots$ の可算個. X を $Po(\mu)$ の確率変数とすると,

$$P(X = x) = e^{-\mu} \frac{\mu^x}{x!}$$

で与えられる. $Po(\mu)$ の期待値と分散は μ となる. (練習問題)

4 連続型確率分布

1. 一様分布 $U(a, b)$ X が区間 $[a, b]$ で一様分布しているとは, 分布密度関数が

$$f(x) = \frac{1}{b-a}$$

で与えられていることを言う. このとき, $a \leq c < d \leq b$ に対して

$$P(c < X \leq d) = \int_c^d \frac{1}{b-a} dx = \frac{d-c}{b-a}$$

となる．平均は

$$E(X) = \int_a^b \frac{x}{b-a} dx = \frac{1}{b-a} \left[\frac{x^2}{2} \right]_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2}$$

と計算する．同様にして $V(X) = \frac{(b-a)^2}{12}$ となる．

2. 指数分布 $Ex(\lambda)$ 非負の値をとる確率変数 X がパラメータ λ の指数分布しているとはその分布密度関数が $[0, \infty)$ 上で

$$f(x) = \lambda e^{-\lambda x}$$

となる時に言う．平均は部分積分をして得られる．

$$\begin{aligned} E(X) &= \int_0^{\infty} x \lambda e^{-\lambda x} dx \\ &= [-x e^{-\lambda x}]_0^{\infty} + \int_0^{\infty} e^{-\lambda x} dx \\ &= \frac{1}{\lambda}. \end{aligned}$$

同様にして $V(X) = E(X^2) - (E(X))^2 = \frac{1}{\lambda^2}$ となる．

3. 正規分布 $N(\mu, \sigma^2)$ 二つのパラメータを持つ．確率変数 X が正規分布 $N(\mu, \sigma^2)$ にしたがうとはその分布密度関数が

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

となるときに言う． $\sigma \geq 0$ とする．次の式（微分積分学 2 で習う）を使うと正規分布 $N(\mu, \sigma^2)$ の平均と分散が計算できる．

$$\int_{-\infty}^{\infty} e^{-x^2/2} dx = \sqrt{2\pi}$$

確かめてみよう． X の分布が $N(\mu, \sigma^2)$ とすると

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\} dx \\ &= \int_{-\infty}^{\infty} (\sigma y + \mu) \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{y^2}{2} \right\} \sigma dy \quad (y = \frac{x-\mu}{\sigma}) \\ &= 0 + \mu \frac{\sqrt{2\pi}}{\sqrt{2\pi}} \\ &= \mu, \end{aligned}$$

$$\begin{aligned}
V(X) &= \int_{-\infty}^{\infty} (x - \mu)^2 \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x - \mu)^2}{2\sigma^2}\right\} dx \\
&= \int_{-\infty}^{\infty} \sigma^2 y^2 \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{y^2}{2}\right\} dy \quad (y = \frac{x-\mu}{\sigma}) \\
&= \left[-\sigma^2 y \exp\left\{-\frac{y^2}{2}\right\}\right]_{-\infty}^{\infty} + \sigma^2 \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{y^2}{2}\right\} dy \\
&= \sigma^2.
\end{aligned}$$

平均 0 分散 1 の正規分布 $N(0, 1)$ を標準正規分布という。さっきの変数変換がほぼ示しているが、次の定理が知られている。

定理 4.1 X が正規分布に従うとき、その一次変換 $Z = aX + b$ もまた正規分布になる。

$N(\mu, \sigma^2)$ に従う確率変数 X に対して $Z = (X - \mu)/\sigma$ と標準化すると、前回やったことにより Z は平均 0、分散 1 となっているので、上の定理から Z は $N(0, 1)$ に従う。

例 4.1 Z が標準正規分布に従うとき $P(Z \geq a) = 0.08$ となる a を標準正規分布表から求めてみる。

$$P(Z \geq a) = P(0 \leq Z < \infty) - P(0 < Z < a) = 0.5 - P(0 < Z < a)$$

だから $P(0 < Z < a) = 0.42$ を満たす a を正規分布表で探す。
 $P(0 \leq Z \leq 1.40) = 0.4192$ で $P(0 \leq Z \leq 1.41) = 0.4207$ だから a は 1.40 と 1.41 の間。線形に比例配分して

$$a = 1.40 + 0.01 \times \frac{8}{8 + 7} = 1.405$$

と計算する。