数理統計学まとめ（その３）：第 3 章 確率変数と確率分布

1 確率変数

サイコロを投げるとき：結果を見るまでは出る目は何かわからない．X_1, X_2, X_3 をそれぞれ 1 回目から 3 回目までのサイコロを振ったときの出る目の数とすると，これらが何かは

- 振ってみるまでわからない．
- もう一度 3 回振ったときも何が出るかはやはりわからない．
- それぞれの目が出る「確率」は 1/6 とわかっている．

確率変数：取る値の範囲と取る確率だけがわかっている「変数」

1. **確率変数の種類**

   (a) **連続確率変数**：結果の数値が隙間なく現れ得るとき．
   身長，体重，ガソリンの消費量，到着時刻等々

   (b) **離散確率変数**: 結果の数値がとびとびで並べられるとき
   またはいくつかの結果が出てくるが，それをとびとびの数字に表せるとき．
   サイコロの目，1 分間にある地点を通る自動車の数，世論調査の結果（「賛成」「反対」を 0 と 1 に表す）等々．可算無限個までとる可能性がある．

2. **確率変数の分布**

   (a) **離散確率変数の分布**：X が取り得る値の全体が有限個のときこれらを $x_1, \ldots, x_n$ と書く．可算無限個のときは $x_1, x_2, \ldots$ と書く．それぞれの x_i に対して

   $$P(X = x_i) = p(x_i) \quad i = 1, 2, \ldots, n.$$

   で **確率変数 X の分布**が特徴づけられる．

   $$P(a < X \leq b) = \sum_{i: a < x_i \leq b} p(x_i)$$

   となる．離散確率変数の分布とはそれぞれの値をとる確率の表を与えることと同じである．

例 1.1 サイコロを 2 回振るとき，1 回目と 2 回目のでた目の和を Y とするとき，Y の分布は次の表で与えられる．

Y	2	3	4	5	6	7	8	9	10	11	12
確率	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

(b) **連続確率変数の分布**：$-\infty < a \leq b < \infty$ に対して

$$P(a < X \leq b) = \big(X \text{ が } a \text{ より大きく } b \text{ 以下となる確率}\big)$$

と与える．応用上多くの場合は，ある関数 $f(x)$ があって

$$P(a < X \leq b) = \int_a^b f(x)\,dx$$

と書ける．この $f(x)$ を確率変数 X の**分布密度関数**という．

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

である．

3. 分布関数：確率変数 X に対して実数 t の関数

$$F(t) = F_X(t) = P(X \leq t)$$

を確率変数 X の**分布関数**と呼ぶ．

注意 1.1 分布関数が与えられると分布は一つに決まる事が知られている．

2 平均と分散

確率変数 X は分布を持っているので，その分布の特性値である**平均** $\mu = E(X)$ や**分散** $\sigma^2 = V(X)$ を考えることができる．

1. 平均 μ：

(a) 離散確率変数のとき

$$\mu = E(X) = \sum_i x_i P(X = x_i)$$

とする．ただし，右辺の和は X のとりうる値全体についてとる．つまり X が有限個の $x_1, \ldots, x_n$ のどれかの値をとるときは $\sum_i = \sum_{i=1}^n$ と理解し，X が無限個の値 $x_1, x_2, \ldots$ をとるときは $\sum_i = \sum_{i=1}^\infty$ と読む．

(b) 連続型確率変数のとき，密度関数 $f(x)$ を持つ場合は

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

と書ける．以下，連続型のときは密度関数がある場合のみ考える．

2. 分散：

(a) 離散確率変数のとき

$$\sigma^2 = V(X) = \sum_i (x_i - \mu)^2 p(x_i)$$

(b) 連続型確率変数のとき

$$\sigma^2 = V(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x)\, dx$$

で与えられる．

関数 $G(x)$ の中に確率変数 X を代入した $Y = G(X)$ は新しい確率変数になるが，この平均を以下で定義する．

$$E(G(X)) = \sum_i G(x_i) P(x_i), \qquad \text{(離散型)}$$

$$E(G(X)) = \int_{-\infty}^{\infty} G(x) f(x)\, dx. \qquad \text{(連続型)}$$

$V(X)$ は $G(x) = (x - \mu)^2$ に対する $E(G(X)) = E((X - \mu)^2)$ を考えていることになる．

(c) **分散公式**

$$V(X) = E(X^2) - \mu^2$$

証明　離散確率変数の場合に証明しておく.

$$\begin{aligned} V(X) &= E((X-\mu)^2) = E(X^2 - 2\mu X + \mu^2) \\ &= \sum_i (x_i^2 - 2\mu x_i + \mu^2) p(x_i) \\ &= \sum_i x_i^2 p(x_i) - 2\mu \sum_i x_i p(x_i) + \mu^2 \sum_i p(x_i) \\ &= E(X^2) - 2\mu E(X) + \mu^2 = E(X^2) - \mu^2. \end{aligned}$$

上で $\sum_i p(x_i) = 1$ と $E(X) = \mu$ であることを使った.

定理 2.1 (教科書 p.55, 定理 3.1)

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

証明　定義にしたがって離散確率変数の場合に証明する.

$$E(aX+b) = \sum_i (ax_i+b)p(x_i) = \sum_i ax_i p(x_i) + \sum_i bp(x_i) = aE(X)+b.$$

$$\begin{aligned} V\big(aX+b\big) &= E\big(\{aX+b-(aE(X)+b)\}^2\big) \\ &= E\big(\{a(X-E(X))\}^2\big) \\ &= E\big(a^2(X-E(X))^2\big) = a^2 E\big((X-E(X)^2\big) = a^2 V(X). \end{aligned}$$

3. **標準偏差**：$\sigma^2 = V(X)$ の正の平方根 $\sqrt{\sigma^2}$ を σ と書き，X の**標準偏差**と呼ぶ.

   標準化 確率変数 X について，$E(X) = \mu, V(X) = \sigma^2$ のとき，

   $$Z = \frac{X-\mu}{\sigma}$$

   を X の**標準化**と呼ぶ．標準化した確率変数は平均が 0, 分散が 1 になっている.