第5回：確率変数と確率分布

尚　晋
大学院経済学研究科　助教

2025年5月20日

今日のポイント

1. 確率変数と確率分布
2. 確率変数の期待値と分散

1　確率変数と確率分布

1.1　離散型の確率変数の確率分布

離散型確率変数　とりうる各値に対しそれぞれ確率が与えられている変数を**確率変数**という．確率変数は，Xのように大文字を用いて表す．例えば、サイコロを振って出る目Xは確率変数である．正しいさいころなら、下記が成り立つ。さらに、2個のさいころを同時に振って出る目X_1, X_2の和X_1+X_2も確率変数で，各値の確率は表5.3, 図5.3となる。

$$P(X=1)=1/6, P(X=2)=1/6, ..., P(X=6)=1/6 \tag{1a}$$
$$\text{一般には，} P(X=1)=p_1, P(X=2)=p_2, ..., P(X=6)=p_6 \tag{1b}$$
$$\text{ここで，} p_1 \geq 0, p_2 \geq 0, ..., p_6 \geq 0, p_1+p_2+...+p_6=1 \text{が成り立つ．} \tag{1c}$$

離散型確率分布　一般に可算集合$\{x_1, x_2, ...\}$の中の値をとる確率変数Xは**離散型**と言われ，それぞれ値の確率：

$$P(X=x_k)=f(x_k) \quad (k=1,2,...) \tag{2}$$

をXの**確率分布(probability distribution)**という，ただし，fは

$$f(x_k) \geq 0, \quad (k=1,2,...)\text{かつ}\sum_{k=1}^{\infty}=1 \tag{3}$$

の条件を満たす．このfを**離散型の確率分布**という．

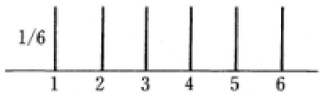


図5.2 さいころの目の確率
どの目も等確率で出現する．離散型一様分布と呼ばれる．

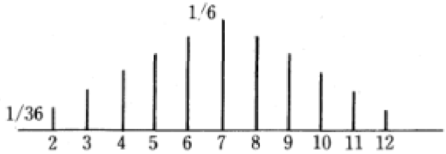


図5.3 2個のさいころの目の和の確率
和の確率分布に峰(モード)が出現する．これは中心極限定理の初期段階である(第8章参照)．

表5.3 2個のさいころの目の和の確率の表

和	2	3	4	5	6	7	8	9	10	11	12
(和/2)	1	3/2	2	5/2	3	7/2	4	9/2	5	11/2	6
確　率	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

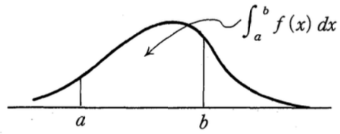


図5.4 密度関数の図
密度関数は確率論の重要な概念だが，目でもわかりやすい．曲線下の面積がその区間の確率となる．$f(x)$ が 0 である区間の確率は 0 である．

1.2 連続型の確率変数の確率分布

連続型確率変数 厳密に測定すると実現値が無限に存在するため、それぞれの実現値に番号が振れない確率変数のこと．

連続型確率変数の確率分布 確率変数Xのとる値が関数$f(x)$によって，

$$P(a \leq X \leq b) = \int_a^b f(x)dx \tag{4}$$

と表される場合，Xは連続型の確率分布を持つという(図5.4)．ただし，全てのxに対して，

$$f(x) \geq 0, \quad \text{かつ} \quad \int_{-\infty}^{\infty} f(x)dx = 1 \tag{5}$$

を満たす．関数$f(x)$をXの**確率密度関数**(probability density function)あるいは単に**密度関数**という．確率変数Xが確率分布$f(x)$をもつとき，Xは$f(x)$に「従う」という．

密度関数(連続型確率分布)の性質:

- aとbががきわめて近いとすると，
$$P(x \leq X \leq x + \Delta x) \fallingdotseq f(x) \cdot \Delta x \tag{6}$$
である．「確率密度」とは，確率を求めるために長さΔxに乗ずるべき「密度」という意味で，$f(x)$が大きい所には確率が濃くあるということを示している．
- 連続型の確率分布においては，一点の確率は0となる．なぜなら，$a = b$とおけば，
$$P(X = a) = 0 \tag{7}$$
となるから．
- 上記の性質から，$P(a \leq X \leq b) = P(a < X < b) = P(a < X \leq b) = P(a \leq X < b)$.

1.3　累積分布関数

ある値以下の確率が必要なことがある．たとえば，飛行機の安全性を考えるとき，現実の積載重量Xが最大安全値を下回る確率は，この飛行機の安全確率となる．

累積分布関数 確率Xに対して，xを実数とするときx以下の確率

$$F(x) = P(X \leq x) \tag{8}$$

をXの積分布関数(cumulative distribution function)と呼ぶ．

連続型の場合 連続型の場合は$F(x)$はfの定積分となる．

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)du \tag{9}$$

微分積分学の基本定理によって，$F(x)$から$F'(x) = f(x)$として元の密度関数$f(x)$が復元される．従って，$F(x)$でも$f(x)$でも確率分布を表すことができる．

離散型の場合 離散型の場合は確率を$-\infty$からはxまで積み上げたものになる：

$$F(x) = P(X \leq x) = \sum_{u \leq x} f(u) \tag{10}$$

累積分布関数の性質:

- 単調増加関数：
$$x_1 < x_2 \quad \text{ならば} \quad F(x_1) \leq F(x_2) \tag{11}$$
- 範囲：
$$\begin{aligned} x \to \infty \text{の時} \quad F(x) \to 1 \\ x \to -\infty \text{の時} \quad F(x) \to 0 \end{aligned} \tag{12}$$
- 右連続(理解するには表5.4と図5.9を参照)：
$$\text{各点}x\text{で}\epsilon \downarrow 0\text{の時} \quad F(x+\epsilon) \to F(x) \tag{13}$$

表 5. 4

x	$f(x)$	$F(x)$
2	1/36	1/36＝0. 028
3	2/36	3/36＝0. 083
4	3/36	6/36＝0. 167
5	4/36	10/36＝0. 278
6	5/36	15/36＝0. 417
7	6/36	21/36＝0. 583
8	5/36	26/36＝0. 722
9	4/36	30/36＝0. 833
10	3/36	33/36＝0. 917
11	2/36	35/36＝0. 972
12	1/36	1

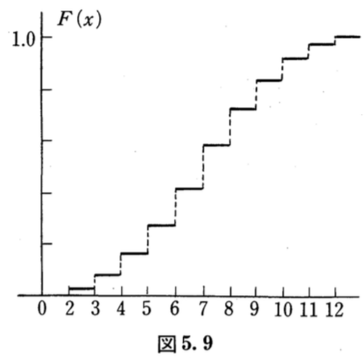


図 5. 9

2 個のさいころの目の和の累積分布関数$F(x)$とその階段状グラフ（表 5. 5，図 5. 9）
2，3，…，12 の各整数点で不連続（ただし右からは連続），それ以外では連続である．増加は不連続点でのジャンプに限る．典型的な離散型の確率分布である．階段の蹴上げの高さが$f(x)$であるが，表の$F(x)$は整数値に対してのみ表示した．

補足 1.1 累積分布関数の右連続性：離散型分布（サイコロ）の場合

■右連続の定義

右連続とは「右から近づいたとき、右側からの極限は関数値と一致する」こと．

関数 $f(x)$ が点 $x=a$ で **右連続（right-continuous）** であるとは、

$$\lim_{x \to a^+} f(x) = f(a)$$

が成り立つことをいう．つまり、a より大きい値（右側）から a に近づけたときの極限が $f(a)$ と等しいことを意味する．

次の2つの極限表記($\lim_{x\to 3^+} F(x)$と$\lim_{x\downarrow 3} F(x)$)は、どちらも「$x$ を3より大きい値から近づけたときの $F(x)$ の極限値」を意味する．

■サイコロの例

サイコロの出目 $X \in \{1,2,3,4,5,6\}$ に対して，確率は一様分布：

$$P(X=k) = \frac{1}{6} \quad (k=1,2,\ldots,6)$$

このときの累積分布関数 $F(x) = P(X \le x)$ は：

$$F(x) = \begin{cases} 0 & (x<1) \\ \frac{1}{6} & (1 \le x < 2) \\ \frac{2}{6} & (2 \le x < 3) \\ \frac{3}{6} & (3 \le x < 4) \\ \frac{4}{6} & (4 \le x < 5) \\ \frac{5}{6} & (5 \le x < 6) \\ 1 & (x \ge 6) \end{cases}$$

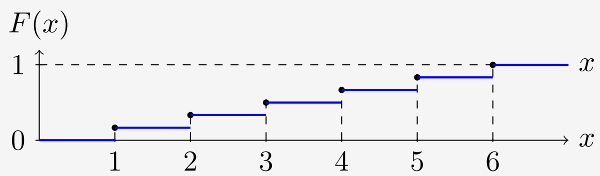
この関数において，例えば $x=2$ における右極限と左極限を考えると：

$$\lim_{x\to 2^+} F(x) = \frac{2}{6} = F(2) \quad （右連続）$$

$$\lim_{x\to 2^-} F(x) = \frac{1}{6} \ne F(2) \quad （左連続ではない）$$

よって，累積分布関数はすべての点で右連続だが，ジャンプのある点では左連続ではない．

図：離散分布（サイコロ）の累積分布関数



2 確率変数の期待値と分散

2.1 期待値

確率変数はいろいろの値をとるが，それらの値を代表する**平均**(正確には，**確率の重みつき平均**)が考えられる．これが**期待値**である．

期待値 確率変数Xに対して，それがとる値の**重みつき平均**を，確率変数の期待値といい，

$E(X)$と書く：

$$E(X) = \begin{cases} \sum_x x f(x) & (\text{離散型}) \\ \int_{-\infty}^{\infty} x f(x)\, dx & (\text{連続型}) \end{cases} \tag{14}$$

x の関数$g(X)$についても同様に期待値を定義する.

$$E(g(X)) = \begin{cases} \sum_x g(x) f(x) & (\text{離散}) \\ \int_{-\infty}^{\infty} g(x) f(x)\, dx & (\text{連続}) \end{cases}$$

例1 確率変数：

$$X := \begin{cases} 1 & \text{with probability } p \\ 0 & \text{with probability } 1-p \end{cases}$$

の期待値：

$$E(X) = \sum_x x f(x) = 1 \cdot p + 0 \cdot (1-p) = p$$

例2 指数分布：

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

の期待値：

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x f(x)\, dx \\ &= \int_0^{\infty} x \lambda e^{-\lambda x}\, dx \\ &= \frac{1}{\lambda} \end{aligned}$$

期待値の演算Eの性質

$$(a) E(c) = c \qquad (\text{cは定数}) \tag{15a}$$
$$(b) E(X+c) = E(X) + c \tag{15b}$$
$$(c) E(cX) = cE(X) \tag{15c}$$
$$(d) E(X+Y) = E(X) + E(Y) \qquad (\text{期待値の加法性}) \tag{15d}$$

性質(d)について、2個のさいころの目の和の確率分布は表5．3でもとめてあるので，$E(X_1 + X_2) = 2 \cdot (1/36) + 3 \cdot (1/36) + ... = 252/36 = 7$，他方，$E(X_1) = 7/2, E(X_2) = 7/2$より，たしかに$(d)$が成立している.

2.2　分散と標準偏差

Xを1個のさいころを振ったとき出た目，Yを2個のさいころを振ったときの2個の目X_1, X_2の相加平均を $Y = (X_1 + X_2)/2$とすると，それぞれの期待値は

$$\begin{aligned} E(X) &= 7/2 \\ E(Y) &= E\{(X_1 + X_2)/2\} \\ &= \{E(X_1) + E(X_2)\}/2 = (7/2 + 7/2)/2 = 7/2 \end{aligned}$$

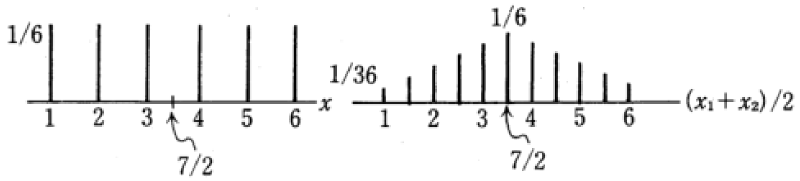


図 5.11 さいころの確率分布の期待値への集中の度合

となり，両者の期待値は等しい．しかし，Xの確率分布は$1, 2, ..., 6$の上に一様にばらついているのに対し，Yの確率分布は期待値$7/2$に集中している(図5.11).

期待値は確率変数の重要な指標ではあるが，期待値が同じでも，異なった確率分布はいくらでもある．いまもう一つの重要な指標が，上に述べた「ばらっき」である．集中やばらっきを表すのに，期待値$E(X)$からのずれの量$X - E(X)$を考える．今後簡単のために，期待値を$\mu = E(X)$で表そう．$X - \mu$はそれ自体確率変数なので，目安としてその平均（期待値）を考えると

$$E(X - \mu) = E(X) - \mu = \mu - \mu = 0$$

となり，ずれ$X - \mu$が正負打ち消しあって0となるのである．かわりに，下記の定義を使う。

分散

$$V(X) = E\{(X - \mu)^2\} \tag{16}$$

と定義しこれをXの**分散**という．上の定義から，必ず，$V(X) \geq 0$ であり，$V(X)$の値が大きいほどXのばらつきは大きい.

分散の計算は，定義から，

$$V(X) = \begin{cases} \sum_x (x - \mu)^2 f(x) & \text{（離散型）} \\ \int_{-\infty}^{\infty} (x - \mu)^2 f(x)\, dx & \text{（連続型）} \end{cases}$$

これを直接使うよりは，下記の等式を用いてもいい。

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

証明：$V(X) = E\{(X - \mu)^2\} = E(X^2) - 2\mu E(X) + \mu^2 = E(X^2) - \mu^2 = E(X^2) - (E(X))^2$

分散の演算Vの性質

$$(a) V(c) = 0 \qquad \text{（定数の変動はゼロ）} \tag{17a}$$

$$(b) V(X + c) = V(X) \tag{17b}$$

$$(c) V(cX) = c^2 V(X) \tag{17c}$$

標準偏差 分散の平方根を**標準偏差**という。

$$D(X) = \sqrt{V(X)} \tag{18}$$

これは分散と同じく「ばらつき」を表すが，分散の定義で二乗したのを戻したので，$D(X)$はもとのXと同じ次元をもつ．分散の値をσ^2で，標準偏差の値をσで表すことが多い.

2.3　標準化

標準化 確率変数の標準化：Xから期待値$E(X)$を引き、標準偏差$\sqrt{V(X)}$で割ることで作られる新たな確率変数

$$Z = \frac{X - E(X)}{\sqrt{V(X)}} \tag{19}$$

を標準化された確率変数と呼ぶ.

標準化された確率変数Zの性質

$$(a) E(Z) = 0 \tag{20a}$$
$$(b) V(Z) = 1 \tag{20b}$$

いかなる確率変数も，その期待値を引き，さらにその標準偏差（分散の平方根）の尺度で割れば，期待値は 0 に，分散は 1 に調整される．この変換を**確率変数の標準化**といい，Zを標準化変数という.

式20aの証明：$E(Z) = E\{\frac{X-E(X)}{\sqrt{V(X)}}\} = E(\frac{X}{\sqrt{V(X)}}) - \frac{E(X)}{\sqrt{V(X)}} = \frac{E(X)}{\sqrt{V(X)}} - \frac{E(X)}{\sqrt{V(X)}} = 0$

式20bの証明：

$$\begin{aligned} V(Z) &= E(Z^2) - (E(Z))^2 \\ &= E(Z^2) \qquad (E(Z)\text{は}0\text{なので}) \\ &= E\{(\frac{X - E(X)}{\sqrt{V(X)}})^2\} \\ &= \frac{1}{V(X)} E\{(X - E(X))^2\} \\ &= \frac{1}{V(X)} V(X) = 1 \end{aligned}$$