

第 158 講 二項分布

学習目標

二項分布が表す状況を判断し、条件を確認して確率を求められるようになる。また、二項分布の期待値・分散・標準偏差を用いて分布の特徴を説明できるようになる。

二項分布が表すもの

一定の試行を繰り返すときの成功回数

この講で扱う装置・種子・くじなどの設定は、すべて数学的理解のために作った架空の教材設定である。

二項分布では、一定の条件で試行を繰り返したとき、注目する事象が何回起こるかを考える。

ここでいう成功とは、注目する事象が起こることを表す呼び名である。失敗はその余事象として分類する。

したがって、「成功」「失敗」は元の試行結果が必ず 2 種類しか存在しないという意味ではなく、注目する事象とそれ以外の事象に分けた分類である。

二項分布として扱うには、固定された n 回の試行、一定の成功確率 p 、全試行の独立性を確認する必要がある。

二項分布の定義

n を正整数、 p を $0 \leq p \leq 1$ を満たす数とする。

成功確率が一定の独立な試行を、最初から決めた n 回だけ行う。各回の成功確率を p とし、成功回数を確率変数 X とする。

このとき、 X が従う分布を二項分布といい、 $X \sim B(n, p)$ と表す。

二項分布として扱うためには、試行回数が固定されていること、各回の成功確率が同じであること、全試行が独立であることが必要である。

二項分布の確率

1 つの並びの確率と並び方の個数

この講では、組合せの数を ${}_m C_r$ と表す。ここで m は非負整数、 r は $0 \leq r \leq m$ を満たす整数であり、 m 個から r 個を選ぶ方法の数を表す。また、0 個から 0 個を選ぶ方法は 1 通りであるため、 ${}_0 C_0 = 1$ と定める。本講では、この記号を成功する位置を選ぶ個数を表すために用いる。

n 回の独立な試行で、ちょうど k 回成功する確率を考える。ただし、ここでは $0 < p < 1$ とし、 n を正整数、 k を整数として、 $0 \leq k \leq n$ とする。

成功する位置を 1 つ固定すると、成功が k 回、失敗が $n - k$ 回であるから、独立性よりその 1 つの並びになる確率は、 $p^k(1 - p)^{n-k}$ である。

一方、成功する位置の選び方は、 ${}_n C_k$ 個ある。

例えば、 $n = 5$ 、 $k = 2$ の場合、成功する位置は 5 個の場所から 2 個を選ぶため、 ${}_5 C_2$ 個の並びがある。

これらの並びは互いに同時には起こらない事象である。そのため、それぞれの並びの確率を足し合わせることで、成功回数が k 回になる確率が求まる。

二項分布の確率公式

$0 < p < 1$ のとき、二項分布 $X \sim B(n, p)$ について、

$P(X = k) = {}_n C_k p^k (1 - p)^{n-k}$ である。

ただし、 $k = 0, 1, \dots, n$ であり、それ以外の整数 k では、 $P(X = k) = 0$ とする。

二項分布で間違えやすい点

例えば、5 回中 2 回成功する場合、特定の 1 つの並びの確率は、

$p^2(1 - p)^3$ であり、これは正しい確率である。

しかし、成功回数が 2 回になる確率を求める場合は、この 1 つの並びだけでは不十分である。成功する位置の選び方が複数あるため、 ${}_5 C_2 p^2 (1 - p)^3$ とする必要がある。

また、各回の成功確率が同じであることだけでは、独立とはいえない。独立とは、ある試行の結果が他の試行の結果へ影響しないことである。

非復元抽出では、前の結果によって残る対象が変化するため、一般に独立性が失われる。ただし、例えば全てのカードが同じ色である場合は、結果が常に一定となるため、別の性質として独立性を確認できる場合がある。

p が 0 または 1 の場合

$p = 0$ または $p = 1$ では、公式へそのまま代入して 0^0 を考えない。

$p = 0$ では必ず失敗するため、 $P(X = 0) = 1$ である。

$p = 1$ では必ず成功するため、 $P(X = n) = 1$ である。

二項分布の期待値・分散・標準偏差

期待値の導出

ここでは $0 < p < 1$ とする。

二項分布 $X \sim B(n, p)$ について、

$$E(X) = \sum_{k=0}^n k {}_n C_k p^k (1 - p)^{n-k}$$

である。

$k = 0$ の項は 0 になるので、 $k \geq 1$ について考える。

組合せの恒等式

$$k {}_n C_k = n {}_{n-1} C_{k-1}$$

この恒等式は組合せの数え方が 2 通りあることから分かる。

$1 \leq k \leq n$ について、 n 個から k 個を選び、その中の 1 個へ印を付けると考える。この数え方が、 $k {}_n C_k$ である。

一方、先に n 個の中から印を付ける 1 個を選び、残りの $n - 1$ 個から印を付けたものと合わせて k 個になるように $k - 1$ 個を選ぶ。この数え方が、 $n {}_{n-1} C_{k-1}$ である。

同じ対象を数えているため、 $k {}_n C_k = n {}_{n-1} C_{k-1}$

が成り立つ。

これを用いると、

$$\begin{aligned}
E(X) &= \sum_{k=1}^n n {}_{n-1}C_{k-1} p^k (1-p)^{n-k} \\
&= np \sum_{k=1}^n {}_{n-1}C_{k-1} p^{k-1} (1-p)^{n-k}
\end{aligned}$$

となる。

ここで $j = k - 1$ とおくと、

$$\begin{aligned}
E(X) &= np \sum_{j=0}^{n-1} {}_{n-1}C_j p^j (1-p)^{n-1-j} \\
&= np \{p + (1-p)\}^{n-1} \\
&= np
\end{aligned}$$

となる。

分散の導出

ここでは $n \geq 2$ 、 $0 < p < 1$ とする。

まず、 $E(X(X-1))$ を考える。

$$E(X(X-1)) = \sum_{k=0}^n k(k-1) {}_nC_k p^k (1-p)^{n-k}$$

である。

$k = 0, 1$ の項は 0 になるので、 $k \geq 2$ について考える。

組合せの恒等式

$$k(k-1) {}_nC_k = n(n-1) {}_{n-2}C_{k-2}$$

この恒等式も組合せの数え方が 2 通りあることから分かる。

$2 \leq k \leq n$ について、 n 個から k 個を選び、その中から順序を区別した 2 個へ印を付けると考える。

先に k 個を選んでから印を付けると、 $k(k-1) {}_nC_k$ となる。

一方、先に印を付ける順序付きの 2 個を選び、残りの $n-2$ 個から $k-2$ 個を選ぶと、

$$n(n-1) {}_{n-2}C_{k-2}$$

となる。

同じ対象を数えているため、 $k(k-1) {}_nC_k = n(n-1) {}_{n-2}C_{k-2}$

が成り立つ。

これを用いると、

$$\begin{aligned}
E(X(X-1)) &= \sum_{k=2}^n n(n-1) {}_{n-2}C_{k-2} p^k (1-p)^{n-k} \\
&= n(n-1)p^2 \sum_{k=2}^n {}_{n-2}C_{k-2} p^{k-2} (1-p)^{n-k}
\end{aligned}$$

となる。

$j = k - 2$ とおくと、

$$\begin{aligned}
E(X(X-1)) &= n(n-1)p^2 \sum_{j=0}^{n-2} {}_{n-2}C_j p^j (1-p)^{n-2-j} \\
&= n(n-1)p^2 \{p + (1-p)\}^{n-2} \\
&= n(n-1)p^2
\end{aligned}$$

となる。

ここで、 $X^2 = X(X-1) + X$ であるから、各値について足し合わせると、

$$\begin{aligned}
E(X^2) &= E(X(X-1)) + E(X) \\
&= n(n-1)p^2 + np
\end{aligned}$$

となる。

したがって、

$$\begin{aligned}
V(X) &= E(X^2) - \{E(X)\}^2 \\
&= n(n-1)p^2 + np - n^2p^2 \\
&= np(1-p)
\end{aligned}$$

となる。

$n=1$ の場合は直接計算する。

$$E(X) = p$$

$$E(X^2) = p \text{ より、 } V(X) = p - p^2 = p(1-p) \text{ となる。}$$

また、 $p=0$ の場合は常に $X=0$ であるため、 $E(X)=0$ 、 $V(X)=0$ である。

$p=1$ の場合は常に $X=n$ であるため、 $E(X)=n$ 、 $V(X)=0$ である。

二項分布の基本性質

n を正整数、 $0 \leq p \leq 1$ とする。

$X \sim B(n, p)$ のとき、 $E(X) = np$

$$V(X) = np(1-p)$$

$\sigma = \sqrt{np(1-p)}$ である。

標準偏差 σ は非負の値として定める。

例題

例題 1

公平なコインを独立に3回投げる。表が出る回数を確率変数 X とする。

(1) X の確率分布を求めよ。

(2) $E(X)$ 、 $V(X)$ 、標準偏差 σ を求めよ。

例題 1 の解答・解説

表が出る確率は $p = \frac{1}{2}$ であり、3回の独立な試行であるから、

$$X \sim B\left(3, \frac{1}{2}\right)$$

である。

したがって、

$$P(X = k) = {}_3C_k \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{3-k}$$

より、

k	0	1	2	3
$P(X = k)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

となる。

期待値は、

$$E(X) = np = 3 \times \frac{1}{2} = \frac{3}{2}$$

である。

分散は、 $V(X) = np(1-p)$ より、

$$V(X) = 3 \times \frac{1}{2} \times \frac{1}{2} = \frac{3}{4}$$

となる。

標準偏差は、

$$\sigma = \sqrt{\frac{3}{4}} = \frac{\sqrt{3}}{2}$$

である。

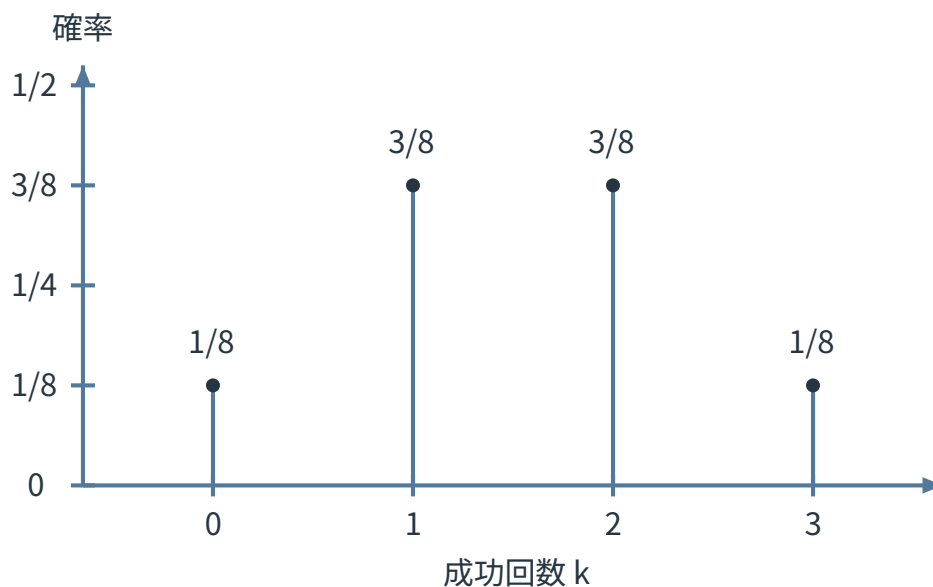


図1 公平なコインを独立に3回投げたときの表の回数 X の確率分布である。横軸の k は成功回数、各棒の高さは $P(X=k)$ を表す。確率は0、1、2、3の4値に対応し、棒の間の面積を確率として読む図ではない。

例題 2

成功確率が $\frac{1}{3}$ である独立な試行を 5 回行う。成功回数を X とする。

(1) ちょうど 2 回成功する確率を求めよ。

(2) $E(X)$ と $V(X)$ を求めよ。

例題 2 の解答・解説

成功確率は $p = \frac{1}{3}$ 、試行回数は $n = 5$ なので、

$$X \sim B\left(5, \frac{1}{3}\right)$$

である。

(1)

$$\begin{aligned} P(X = 2) &= {}_5C_2 \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^3 \\ &= 10 \times \frac{1}{9} \times \frac{8}{27} \\ &= \frac{80}{243} \end{aligned}$$

(2)

$$E(X) = np = 5 \times \frac{1}{3} = \frac{5}{3}$$

である。

また、

$$\begin{aligned} V(X) &= np(1 - p) \\ &= 5 \times \frac{1}{3} \times \frac{2}{3} \\ &= \frac{10}{9} \end{aligned}$$

となる。

例題 3

成功確率が $\frac{1}{4}$ である独立な試行を 4 回行う。成功回数を確率変数 X とする。

1 回以上成功する確率を求めよ。

例題 3 の解答・解説

「1 回以上成功する」は、成功回数 X が 1 以上になることを表す。
直接求めるより、余事象である「1 回も成功しない」確率を考える。
成功しない確率は、

$$1 - \frac{1}{4} = \frac{3}{4}$$

である。
したがって、

$$\begin{aligned} P(X \geq 1) &= 1 - P(X = 0) \\ &= 1 - \left(\frac{3}{4}\right)^4 \\ &= 1 - \frac{81}{256} \\ &= \frac{175}{256} \end{aligned}$$

となる。

演習問題

練習問題 1

ある検査装置は、1 回の検査で製品を「合格」と判定する確率が $\frac{2}{5}$ である。検査は 5 回行い、各検査結果は互いに独立である。合格回数を確率変数 X とする。

- (1) X が従う分布を答えよ。
- (2) ちょうど 3 回合格する確率を求めよ。

練習問題 1 の解答・解説

1 回ごとの合格確率が一定で、試行回数が 5 回に固定され、各回が独立している。
したがって、

$$X \sim B\left(5, \frac{2}{5}\right)$$

である。
(2) ちょうど 3 回合格する確率は、二項分布の公式を使う。

$$\begin{aligned} P(X = 3) &= {}_5C_3 \left(\frac{2}{5}\right)^3 \left(1 - \frac{2}{5}\right)^{5-3} \\ &= 10 \times \frac{8}{125} \times \left(\frac{3}{5}\right)^2 \\ &= \frac{720}{3125} \end{aligned}$$

となる。
ポイントは、特定の 1 つの並びの確率だけではなく、その並び方の個数 ${}_5C_3$ を掛けることである。

練習問題 2

あるゲームを 10 回行う。各回で得点条件を満たす確率は $\frac{3}{10}$ であり、各回の結果は独立である。条件を満たした回数を確率変数 X とする。 $E(X)$ 、 $V(X)$ 、標準偏差 σ を求めよ。

練習問題 2 の解答・解説

成功確率は $p = \frac{3}{10}$ 、試行回数は $n = 10$ なので、

$$X \sim B\left(10, \frac{3}{10}\right)$$

である。

期待値は、 $E(X) = np$ より、

$$E(X) = 10 \times \frac{3}{10} = 3$$

である。

分散は、

$$\begin{aligned} V(X) &= np(1-p) \\ &= 10 \times \frac{3}{10} \times \frac{7}{10} \\ &= \frac{21}{10} \end{aligned}$$

となる。

標準偏差は、

$$\sigma = \sqrt{\frac{21}{10}}$$

である。

練習問題 3

成功確率が $\frac{1}{5}$ である独立な試行を 6 回行う。成功回数を確率変数 X とする。少なくとも 1 回成功する確率を求めよ。

練習問題 3 の解答・解説

少なくとも 1 回成功する確率を直接求めるより、余事象を考える。
余事象は「1 回も成功しない」である。
成功しない確率は、

$$1 - \frac{1}{5} = \frac{4}{5}$$

なので、

$$\begin{aligned} P(X \geq 1) &= 1 - P(X = 0) \\ &= 1 - \left(\frac{4}{5}\right)^6 \\ &= 1 - \frac{4096}{15625} \\ &= \frac{11529}{15625} \end{aligned}$$

となる。

練習問題 4

1 回だけ試行を行い、成功確率を p とする。成功回数を確率変数 X とする。

(1) $0 < p < 1$ のとき、 $P(X = 0)$ と $P(X = 1)$ を求めよ。

(2) $E(X)$ と $V(X)$ を求め、二項分布の公式と一致することを確認せよ。

練習問題 4 の解答・解説

(1) 1 回だけ試行するので、
成功する場合は $X = 1$ 、失敗する場合は $X = 0$ である。

したがって、 $P(X = 1) = p$

$P(X = 0) = 1 - p$ である。

(2) 期待値は、

$$\begin{aligned} E(X) &= 0(1 - p) + 1 \cdot p \\ &= p \end{aligned}$$

となる。

また、 $E(X^2) = 0^2(1 - p) + 1^2p = p$ より、

$$\begin{aligned} V(X) &= E(X^2) - \{E(X)\}^2 \\ &= p - p^2 \\ &= p(1 - p) \end{aligned}$$

となる。

二項分布の公式で $n = 1$ とすると、 $E(X) = 1p = p$

$V(X) = 1p(1 - p) = p(1 - p)$ となり一致する。

練習問題 5

成功回数を確率変数 X とする。

(1) 3 回の独立な試行で成功確率 $p = 0$ の場合の確率分布を求めよ。

(2) 3 回の独立な試行で成功確率 $p = 1$ の場合の確率分布を求めよ。

ただし、主公式へ代入して 0^0 を考えるのではなく、試行の意味から考えよ。

練習問題 5 の解答・解説

(1)

$p = 0$ では成功することはない。

したがって、

k	0	1	2	3
$P(X = k)$	1	0	0	0

である。

(2)

$p = 1$ では必ず成功する。

したがって、

k	0	1	2	3
$P(X = k)$	0	0	0	1

である。

$p = 0, 1$ では主公式の条件 $0 < p < 1$ から外れるため、試行の結果から直接考える。

練習問題 6

赤いカード 1 枚と青いカード 1 枚がある。各回、残ったカードから等確率で 1 枚を引き、引いたカードは戻さずに 2 回順に引く。

A を「1 回目に赤を引く事象」、 B を「2 回目に赤を引く事象」とする。

(1) $P(A)$ 、 $P(B)$ を求めよ。

(2) この 2 つの事象が独立ではないことを説明せよ。

(3) 赤を引いた回数を S とするとき、 $B(2, \frac{1}{2})$ を適用できない理由を説明せよ。

練習問題 6 の解答・解説

C を、1 回目と 2 回目の両方で赤を引く事象とする。

(1) 1 回目は赤と青の 2 枚から 1 枚を選ぶので、

$$P(A) = \frac{1}{2}$$

である。

2 回目に赤となるのは、1 回目に青を引いた場合である。1 回目に青を引く確率は $\frac{1}{2}$ であり、その場合は残ったカードが赤 1 枚になるため、

$$P(B) = \frac{1}{2}$$

となる。

(2) 独立であるなら、 $P(C) = P(A)P(B)$

でなければならない。

しかし、赤いカードは 1 枚しかないため、1 回目と 2 回目の両方で赤を引くことはできない。
したがって、 $P(C) = 0$ であり、

$$P(A)P(B) = \frac{1}{4}$$

だから、 $P(C) \neq P(A)P(B)$ となる。

よって、 A と B は独立ではない。

(3) 各回の赤になる確率が同じでも、独立性は別に確認する必要がある。

この試行では赤の枚数が必ず 1 枚であり、2 回の結果は互いに影響する。

そのため、赤を引いた回数 S について、

$$S \sim B\left(2, \frac{1}{2}\right)$$

とはできない。

練習問題 7

次の 2 つの試行について考える。

- ・ A：当たりが出る確率が $\frac{1}{3}$ のくじを、毎回戻しながら 10 回引く。各回の結果は独立であり、当たり回数を確率変数 X とする。
 - ・ B：当たりが出る確率が $\frac{1}{3}$ のくじを、毎回戻しながら当たりが出るまで引く。各回の結果は独立であり、最初の当たりまでに引く回数を確率変数 N とする。
- それぞれ二項分布として扱えるか説明せよ。

練習問題 7 の解答・解説

A では、試行回数が 10 回に固定され、各回の成功確率が $\frac{1}{3}$ で一定であり、独立である。
したがって、

$$X \sim B\left(10, \frac{1}{3}\right)$$

として扱える。

一方、B では成功するまで試行を続けるため、試行回数が最初から固定されていない。

また、 N は固定された n 回の中で成功した回数ではなく、最初の成功までに必要な試行回数を表す。
したがって、 N には固定回数の二項分布をそのまま適用できない。

練習問題 8

ある農園では、1 個の種子が発芽する確率が $\frac{3}{4}$ であることが分かっている。10 個の種子について調べる。

ただし、各種子の発芽結果は互いに独立であり、調査する個数は最初から 10 個に固定されている。

発芽した種子の個数を確率変数 X とする。

- (1) X が従う分布を答えよ。
- (2) ちょうど 8 個発芽する確率を求めよ。
- (3) 発芽個数の平均と分散を求めよ。

練習問題 8 の解答・解説

発芽確率は $p = \frac{3}{4}$ 、調査個数は $n = 10$ で固定され、各種子の結果は独立である。
したがって、

$$X \sim B\left(10, \frac{3}{4}\right)$$

である。

(2)

$$\begin{aligned} P(X = 8) &= {}_{10}C_8 \left(\frac{3}{4}\right)^8 \left(\frac{1}{4}\right)^2 \\ &= 45 \times \frac{6561}{65536} \times \frac{1}{16} \\ &= \frac{295245}{1048576} \end{aligned}$$

となる。

(3) 平均は、 $E(X) = np$ より、

$$E(X) = 10 \times \frac{3}{4} = \frac{15}{2}$$

である。

分散は、

$$\begin{aligned} V(X) &= np(1 - p) \\ &= 10 \times \frac{3}{4} \times \frac{1}{4} \\ &= \frac{15}{8} \end{aligned}$$

となる。