

## 第2章

# 確率分布と統計的な推測

### Check!

No.	難易度	1回目	2回目	No.	難易度	1回目	2回目
1	*			9	*		
2	*			10	**		
3	**			11	**		
4	*			12	**		
5	*			13	**		
6	*			14	*		
7	*			15	*		
8	*			16	**		

### 例題

No.	難易度	1回目	2回目	No.	難易度	1回目	2回目
B2.1	**			B2.9	***		
B2.2	**			B2.10	**		
B2.3	**			B2.11	***		
B2.4	**			B2.12	***		
B2.5	***			B2.13	**		
B2.6	***			B2.14	**		
B2.7	**			B2.15	**		
B2.8	**			B2.16	***		

### 練習

No.	難易度	1回目	2回目	No.	難易度	1回目	2回目
B2.1	**			B2.9	***		
B2.2	**			B2.10	**		
B2.3	**			B2.11	***		
B2.4	**			B2.12	***		
B2.5	***			B2.13	**		
B2.6	***			B2.14	**		
B2.7	**			B2.15	**		
B2.8	**			B2.16	***		

### Step Up

No.	難易度	1回目	2回目	No.	難易度	1回目	2回目
1	**			9	*****		
2	***			10	**		
3	***			11	***		
4	**			12	***		
5	***			13	***		
6	***			14	***		
7	**			15	***		
8	***			16	***		

# まとめ 1 確率分布

## 1. 確率変数と確率分布

$X$  が  $x_1, x_2, \dots, x_n$  のいずれかの値をとる変数で、 $X$  が 1 つの値  $x_k$  をとる確率  $P$  が  $p_k$  で定まるとき、 $X$  を確率変数という。

また、 $x_k$  と  $p_k$  の対応関係は右の表のようになり、この対応関係を  $X$  の確率分布という。

$X$	$x_1$	$x_2$	……	$x_n$	計
$P$	$p_1$	$p_2$	……	$p_n$	1

ここで、 $p_k \geq 0$  ( $k=1, 2, \dots, n$ ),

$p_1 + p_2 + \dots + p_n = 1$  である。確率変数  $X$  の値が  $a$  となる確率を  $P(X=a)$ 、 $X$  が  $a$  以上  $b$  以下となる確率を  $P(a \leq X \leq b)$  と表す。

(例) 4 枚の硬貨を同時に投げるとき、表の出る枚数を  $X$  とすると、 $X$  は 0, 1, 2, 3, 4 の値をとる確率変数である。 $X=k$  となる確率は

$$P(X=k) = {}_4C_k \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{4-k} \text{ より、確率分布は下のようになる。}$$

$X$	0	1	2	3	4	計
$P$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$	1

$$\begin{aligned} \text{また、} P(0 \leq X \leq 2) &= \frac{1}{16} + \frac{4}{16} + \frac{6}{16} \\ &= \frac{11}{16} \end{aligned}$$

## 2. 確率変数の平均

確率変数  $X$  が右の表に示された分布に従うとき、

$$x_1 p_1 + x_2 p_2 + \dots + x_n p_n$$

を、確率変数  $X$  の平均または期待値といい、 $E(X)$  で表す。

1 次式の平均

$a, b$  が定数で、 $Y = aX + b$  で表されるとき、 $Y$  も確率変数となる。

このとき、 $Y$  の平均  $E(Y)$  は、 $E(Y) = E(aX + b) = aE(X) + b$  で表される。

$X$	$x_1$	$x_2$	……	$x_n$	計
$P$	$p_1$	$p_2$	……	$p_n$	1

◀  $E(X)$  の  $E$  は  
Expectation (期待値)  
に由来

## 3. 確率変数の分散と標準偏差

右の表のような分布をもつ確率変数  $X$  の平均を  $E(X) = m$  とするとき、

$$E((X-m)^2) = (x_1 - m)^2 p_1 + (x_2 - m)^2 p_2 + \dots + (x_n - m)^2 p_n$$

を確率変数  $X$  の分散といい、 $V(X)$  で表す。

また、 $V(X)$  の正の平方根を  $X$  の標準偏差といい、 $\sigma(X)$  で表す。

$$\text{つまり、} \sigma(X) = \sqrt{V(X)} = \sqrt{E((X-m)^2)}$$

分散と標準偏差の性質

$$V(X) = E(X^2) - \{E(X)\}^2, \quad \sigma(X) = \sqrt{E(X^2) - \{E(X)\}^2} \text{ が成り立つ。}$$

## 1 次式の分散と標準偏差

$a, b$  が定数で,  $Y=aX+b$  のとき,  $Y$  の分散  $V(Y)$  と標準偏差  $\sigma(Y)$  は,

$$V(Y)=V(aX+b)=a^2V(X), \quad \sigma(Y)=\sqrt{V(Y)}=|a|\sigma(X)$$

(証明)  $E(Y)=E(aX+b)=aE(X)+b$  より,

$$Y-E(Y)=(aX+b)-\{aE(X)+b\}=a\{X-E(X)\}$$

$$\text{したがって, } V(Y)=E(\{Y-E(Y)\}^2)=E(a^2\{X-E(X)\}^2)$$

$$=a^2E(\{X-E(X)\}^2)=a^2V(X)$$

$$\text{また, } \sigma(Y)=\sqrt{V(Y)}=\sqrt{a^2V(X)}=|a|\sqrt{V(X)}=|a|\sigma(X)$$

## 4. 独立な確率変数

2つの確率変数  $X, Y$  とそれらがとり得る任意の値  $x_i, y_j$  について, 次の式が成り立つとき,  $X, Y$  は独立であるという.

$$P(X=x_i, Y=y_j)=P(X=x_i)P(Y=y_j)$$

## 5. 確率変数の和の平均と独立な確率変数の平均・分散

2つの確率変数  $X, Y$  の和を  $Z=X+Y$  とすると,  $Z$  も確率変数である.  
確率変数の和の平均 (期待値)

$$E(X+Y)=E(X)+E(Y)$$

一般に,  $a, b$  が定数のとき,  $E(aX+bY)=aE(X)+bE(Y)$

独立な確率変数の積の平均 (期待値)

$$X, Y \text{ が独立のとき, } E(XY)=E(X)E(Y)$$

独立な確率変数の和の分散, 標準偏差

$$X, Y \text{ が独立のとき, } V(X+Y)=V(X)+V(Y)$$

$$\sigma(X+Y)=\sqrt{\{\sigma(X)\}^2+\{\sigma(Y)\}^2}$$

一般に,  $X, Y$  が独立で,  $a, b$  が定数のとき,

$$V(aX+bY)=a^2V(X)+b^2V(Y)$$

## 6. 二項分布

1回の試行で事象  $A$  の起こる確率が  $p$ , 起こらない確率が  $q$  のとき,  $n$  回の試行が独立ならば, 事象  $A$  が  $r$  回起こる確率は,  ${}_nC_rp^rq^{n-r} (p+q=1)$  である. よって,  $n$  回の試行のうち, 事象  $A$  の起こる回数を  $X$  とすると,  $X$  の確率分布は次のようになる.

$X$	0	1	.....	$r$	.....	$n$	計
$P$	${}_nC_0q^n$	${}_nC_1pq^{n-1}$	.....	${}_nC_rp^rq^{n-r}$	.....	${}_nC_np^n$	1

このような確率分布を二項分布といい,  $B(n, p)$  で表す.

また, このとき, 確率変数  $X$  は二項分布  $B(n, p)$  に従うという.

二項分布の平均と標準偏差

二項分布  $B(n, p)$  に従う確率変数  $X$  において,  $q=1-p$  とすると, 平均  $E(X)$  と標準偏差  $\sigma(X)$  は,  $E(X)=np$ ,  $\sigma(X)=\sqrt{npq}$  で表される.

## Check!

\*

- 1 袋の中に数字 1, 2, 3 を記入した玉が, それぞれ, 1 個, 2 個, 3 個の合計 6 個入っている. この袋から 2 個の玉を同時に取り出し, それぞれに記入されている数の和を  $X$  とする.  $X$  の確率分布を求めよ. また, 平均, 分散および標準偏差を求めよ.

\*

- 2 確率変数  $X$  の確率分布が右の表で与えられ, 標準偏差が 2 であるとき, 正の定数  $a$ ,  $b$  の値を求めよ.

$X$	$3a$	$a$	$a-1$	計
$P$	$3b$	$5b$	$2b$	1

\*\*

- 3 (1) 確率変数  $X$  の確率分布が, 右の表のように与えられている. このとき,  $Y=aX+b$  ( $a, b$  は定数) で決まる

$X$	$x_1$	$x_2$	……	$x_n$	計
$P$	$p_1$	$p_2$	……	$p_n$	1

確率変数  $Y$  の平均は  $E(Y)=aE(X)+b$  であることを示せ.

- (2) 赤玉が 5 個, 白玉が 3 個入った箱から同時に 4 個の玉を取り出し, 取り出した赤玉 1 個につき 10 ポイントもらえるゲームを行う. ただし, ゲーム 1 回ごとに参加料 25 ポイントを払うものとする.  
このゲームを 1 回行ったとき, 取り出された赤玉の個数を  $X$  とし, 参加料を差し引いて得るポイント数を  $Y$  として,  $Y$  の平均を求めよ.

\*

- 4 100 円硬貨 1 枚, 50 円硬貨 2 枚を同時に投げるとき, 表の出た硬貨の金額の和の平均と分散を求めよ.

\*

- 5 A の袋には①, ①, ②, ②, ③の 5 枚のカードが, B の袋には④, ⑤, ⑤, ⑥, ⑥, ⑥の 6 枚のカードが入っている. この A, B の袋からカードを 1 枚ずつ取り出すとき, カードに書かれている数の積の平均を求めよ.

\*

- 6 10 円硬貨を 5 枚投げる. 次の確率変数の平均, 分散および標準偏差を求めよ.

- (1) 表の出る枚数  $X$   
(2) 表の出る硬貨の合計金額  $Y$

▶▶ 解答編 p.B2-1

1 確率分布略, 平均  $\frac{14}{3}$ , 分散  $\frac{8}{9}$ , 標準偏差  $\frac{2\sqrt{2}}{3}$

2  $a=2, b=\frac{1}{10}$

3 (1) 略 (2) 0

4 平均 100, 分散 3750

5  $\frac{48}{5}$

6 (1) 平均  $\frac{5}{2}$ , 分散  $\frac{5}{4}$ , 標準偏差  $\frac{\sqrt{5}}{2}$  (2) 平均 25, 分散 125, 標準偏差  $5\sqrt{5}$

## 解説

Commentary

## 「確率変数と平均」

1 確率変数の和の平均  $E(X+Y)=E(X)+E(Y)$ 

袋 A にカード①が 3 枚, ②が 2 枚, 袋 B にカード⑥が 1 枚, ⑧が 4 枚入っているとす。2 つの袋から 1 枚ずつカードを引くとき, 袋 A から引いたカードの数  $X$ , 袋 B から引いたカードの数  $Y$  の確率分布は次の表 1, 2 のようになる。

表 1

$X$	1	2	計
$P$	$\frac{3}{5}$	$\frac{2}{5}$	1

表 2

$Y$	6	8	計
$P$	$\frac{1}{5}$	$\frac{4}{5}$	1

表 3

$X \backslash Y$	6	8	計
1	$p_1$	$p_2$	$p_1 + p_2$
2	$p_3$	$p_4$	$p_3 + p_4$
計	$p_1 + p_3$	$p_2 + p_4$	1

数  $X, Y$  を同時に考えたとき, 確率が表 3 のようになっているとすると, 表 1, 2 より,

$$p_1 + p_2 = \frac{3}{5}, \quad p_3 + p_4 = \frac{2}{5}, \quad p_1 + p_3 = \frac{1}{5}, \quad p_2 + p_4 = \frac{4}{5}$$

このとき,  $X+Y$  の確率分布は右のようになるので,  $X+Y$  の平均は次のように計算できる。

$X+Y$	7	8	9	10	計
$P$	$p_1$	$p_3$	$p_2$	$p_4$	1

$$\begin{aligned} E(X+Y) &= 7p_1 + 8p_3 + 9p_2 + 10p_4 \\ &= (1+6)p_1 + (2+6)p_3 + (1+8)p_2 + (2+8)p_4 \\ &= 1 \times (p_1 + p_2) + 2(p_3 + p_4) + 6(p_1 + p_3) + 8(p_2 + p_4) \\ &= \left(1 \times \frac{3}{5} + 2 \times \frac{2}{5}\right) + \left(6 \times \frac{1}{5} + 8 \times \frac{4}{5}\right) = E(X) + E(Y) \end{aligned}$$

一般に,  $a, b$  が定数のとき,  $E(aX + bY) = aE(X) + bE(Y)$  が成り立つ。

2 独立な確率変数  $X, Y$  の積の平均  $E(XY) = E(X)E(Y)$ 

2 つの独立な試行  $A, B$  について,  $A$  に関する確率変数  $X$  と  $B$  に関する確率変数  $Y$  の確率分布が, それぞれ, 次の表 4, 5 のようであるとする。

表 4

$X$	$x_1$	$x_2$	計
$P$	$p_1$	$p_2$	1

表 5

$Y$	$y_1$	$y_2$	計
$P$	$q_1$	$q_2$	1

表 6

$X \backslash Y$	$y_1$	$y_2$	計
$x_1$	$p_1q_1$	$p_1q_2$	$p_1$
$x_2$	$p_2q_1$	$p_2q_2$	$p_2$
計	$q_1$	$q_2$	1

試行  $A, B$  は独立であるから,  $X$  が  $x_i$  で  $Y$  が  $y_j$  となる確率は  $p_iq_j$  で, 確率は表 6 のようになる。

よって,  $X, Y$  の積  $XY$  の確率分布は,

$XY$	$x_1y_1$	$x_1y_2$	$x_2y_1$	$x_2y_2$	計
$P$	$p_1q_1$	$p_1q_2$	$p_2q_1$	$p_2q_2$	1

右の表のようになる。

したがって, 独立な確率変数  $X, Y$  の積  $XY$  の平均は次のようになる。

$$\begin{aligned} E(XY) &= x_1y_1p_1q_1 + x_1y_2p_1q_2 + x_2y_1p_2q_1 + x_2y_2p_2q_2 \\ &= (x_1p_1 + x_2p_2)(y_1q_1 + y_2q_2) = E(X)E(Y) \end{aligned}$$

## 例題 B2.1 確率分布と確率変数の平均

\*\*\*\*

8枚のカード①, ②, ③, ④, ⑤, ⑥, ⑦, ⑧から無作為に4枚を取り出す。取り出された4枚のカードに記されている数の最小値を $X$ とする。

- (1)  $P(X=3)$ を求めよ。 (2)  $X$ の確率分布を求めよ。  
(3)  $X$ の平均(期待値)を求めよ。

- 考え方** (1)  $P(X=3)$ は最小値が3となる確率のことで、これは、「③を取り出し」かつ、「④, ⑤, ⑥, ⑦, ⑧の5枚から3枚取り出す」ときに起こる。  
(2) 確率変数 $X$ のとり得る値は1, 2, 3, 4, 5である。  
(3) 確率変数 $X$ が $X=x_k$ となる確率が $p_k$ のとき、確率変数 $X$ の平均(期待値)は、 $E(X)=x_1p_1+x_2p_2+\cdots+x_np_n$ である。

- 解答** (1) 8枚のカードから無作為に4枚を取り出す方法は、 ${}_8C_4$ 通りである。

③を取り出し、かつ、④, ⑤, ⑥, ⑦, ⑧のカード5枚から3枚を取り出す方法は、 $1 \times {}_5C_3$ (通り)である。

$$\text{よって、} P(X=3) = \frac{1 \times {}_5C_3}{{}_8C_4} = \frac{10}{70} = \frac{1}{7}$$

- (2) 確率変数 $X$ のとり得る値は、1, 2, 3, 4, 5である。

$$P(X=1) = \frac{1 \times {}_7C_3}{{}_8C_4} = \frac{35}{70}, \quad P(X=2) = \frac{1 \times {}_6C_3}{{}_8C_4} = \frac{20}{70},$$

$$P(X=4) = \frac{1 \times {}_4C_3}{{}_8C_4} = \frac{4}{70}, \quad P(X=5) = \frac{1 \times {}_3C_3}{{}_8C_4} = \frac{1}{70}$$

となるから、最小値 $X$ の確率分布は次のとおりである。

$X$	1	2	3	4	5	計
$P$	$\frac{1}{2}$	$\frac{2}{7}$	$\frac{1}{7}$	$\frac{2}{35}$	$\frac{1}{70}$	1

- (3) 最小値 $X$ の平均は、

$$\begin{aligned} E(X) &= 1 \times \frac{1}{2} + 2 \times \frac{2}{7} + 3 \times \frac{1}{7} + 4 \times \frac{2}{35} + 5 \times \frac{1}{70} \\ &= \frac{9}{5} \end{aligned}$$

③から⑧の6枚のカードから4枚を取り出すとしないように注意

(1)と同様に、その数より大きい数が3つ以上ある数となる。

本書では、解答の $P$ を既約分数で表しているが、平均を求める場合などでは、約分していない形の方が計算が楽な場合もある。

$$E(X) = x_1p_1 + x_2p_2 + \cdots + x_np_n$$

## Focus

確率変数 $X$ が $X=x_k$ となる確率が $p_k$ (つまり、 $P(X=x_k)=p_k$ )のとき、  
確率変数 $X$ の平均(期待値)  $E(X)=x_1p_1+x_2p_2+\cdots+x_np_n$   
ただし、 $p_k \geq 0$  ( $k=1, 2, \dots, n$ )、 $p_1+p_2+\cdots+p_n=1$

## 練習 B2.1

\*\*

数字の書かれた玉①, ②, ③, ④, ⑤がそれぞれ2個ずつ入っている袋の中から3個の玉を取り出すとき、玉に書かれている数の最大値を $X$ とする。

- (1)  $P(X=4)$ を求めよ。 (2)  $X$ の確率分布を求めよ。  
(3)  $X$ の平均(期待値)を求めよ。

## 例題 B2.2 1 次式の確率変数の平均・分散・標準偏差 \*\*\*\*

赤玉 4 個と白玉 3 個が入っている袋から、3 個の玉を同時に取り出すとき、赤玉の個数を  $X$ 、白玉の個数を  $Y$  とする。

- (1)  $X$  の平均  $E(X)$ 、分散  $V(X)$ 、標準偏差  $\sigma(X)$  を求めよ。  
 (2)  $Z=4X+2Y$  とするとき、 $Z$  の平均  $E(Z)$ 、分散  $V(Z)$ 、標準偏差  $\sigma(Z)$  を求めよ。

- 考え方** (1) 確率変数  $X$  のとり得る値は 0, 1, 2, 3 であり、分散  $V(X)$ 、標準偏差  $\sigma(X)$  は、 $V(X)=E(X^2)-\{E(X)\}^2$ 、 $\sigma(X)=\sqrt{V(X)}$  を利用して求める。  
 (2)  $X+Y=3$  より、 $Y=-X+3$  を代入すると、 $Z=2X+6$  ( $X$  の 1 次式) となる。  
 1 次式の平均  $E(aX+b)=aE(X)+b$ 、分散  $V(aX+b)=a^2V(X)$  である。

- 解答** (1) 確率変数  $X$  のとり得る値は、0, 1, 2, 3 である。

$$P(X=0)=\frac{{}_3C_3}{{}_7C_3}=\frac{1}{35}, \quad P(X=1)=\frac{{}_4C_1 \times {}_3C_2}{{}_7C_3}=\frac{12}{35},$$

$$P(X=2)=\frac{{}_4C_2 \times {}_3C_1}{{}_7C_3}=\frac{18}{35}, \quad P(X=3)=\frac{{}_4C_3}{{}_7C_3}=\frac{4}{35}$$

より、確率変数  $X$  の確率分布は右のようになる。

$X$	0	1	2	3	計
$P$	$\frac{1}{35}$	$\frac{12}{35}$	$\frac{18}{35}$	$\frac{4}{35}$	1

$$\text{よって, } E(X)=0 \times \frac{1}{35} + 1 \times \frac{12}{35} + 2 \times \frac{18}{35} + 3 \times \frac{4}{35} = \frac{60}{35} = \frac{12}{7}$$

$$\text{また, } E(X^2)=0^2 \times \frac{1}{35} + 1^2 \times \frac{12}{35} + 2^2 \times \frac{18}{35} + 3^2 \times \frac{4}{35} = \frac{120}{35} = \frac{24}{7} \text{ より,}$$

$$V(X)=E(X^2)-\{E(X)\}^2=\frac{24}{7}-\left(\frac{12}{7}\right)^2=\frac{168-144}{49}=\frac{24}{49}$$

$$\text{よって, } \sigma(X)=\sqrt{V(X)}=\sqrt{\frac{24}{49}}=\frac{2\sqrt{6}}{7}$$

- (2)  $X+Y=3$  より、 $Y=-X+3$  であるから、

$$Z=4X+2Y=4X+2(-X+3)=2X+6$$

$$\text{よって, } E(Z)=E(2X+6)=2E(X)+6$$

$$=2 \times \frac{12}{7} + 6 = \frac{66}{7}$$

$$V(Z)=V(2X+6)=2^2V(X)=4 \times \frac{24}{49} = \frac{96}{49}$$

$$\sigma(Z)=\sigma(2X+6)=|2|\sigma(X)=2 \times \frac{2\sqrt{6}}{7} = \frac{4\sqrt{6}}{7}$$

$X, Y$  は独立ではないから、例題 B2.3 で学習する公式

$$\begin{aligned} V(aX+bY) &= a^2V(X)+b^2V(Y) \\ &\text{は利用できない。} \end{aligned}$$

$$\sigma(Z)=\sqrt{V(Z)} \text{ でもよい。}$$

## Focus

$X$  は確率変数、 $a, b$  は定数のとき、 $aX+b$  (1 次式) の平均  $E(aX+b)=aE(X)+b$ 、分散  $V(aX+b)=a^2V(X)$ 、標準偏差  $\sigma(aX+b)=\sqrt{V(aX+b)}=\sqrt{a^2V(X)}=|a|\sigma(X)$

## 練習 B2.2

\*\*\*

例題 B2.2 において、求めた  $X$  の平均  $E(X)$ 、分散  $V(X)$ 、標準偏差  $\sigma(X)$  を用いて、 $Y$  の平均  $E(Y)$ 、分散  $V(Y)$ 、標準偏差  $\sigma(Y)$  を求めよ。  $\rightarrow p.B2-13$  ①~③⑤

## 例題 B2.3 確率変数の和の平均・分散

\*\*\*\*

100 円硬貨 3 枚, 10 円硬貨 4 枚を同時に投げ, 100 円硬貨の表が出た枚数を  $X$ , 10 円硬貨の表が出た枚数を  $Y$  とする. 表が出た硬貨がもらえるものとするとき, もらえる金額  $Z$  の平均と分散を求めよ.

**考え方**  $X$  と  $Y$  が独立でなくても, 和の平均  $E(aX + bY) = aE(X) + bE(Y)$  は成り立つ.  $X$  と  $Y$  が独立のとき, 和の分散  $V(aX + bY) = a^2V(X) + b^2V(Y)$  が成り立つ.

**解答** もらえる金額は,  $Z = 100X + 10Y$  である.

確率変数  $X$  のとり得る値は 0, 1, 2, 3 で,

$$P(X=k) = {}_3C_k \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{3-k} \quad (k=0, 1, 2, 3)$$

確率変数  $Y$  のとり得る値は 0, 1, 2, 3, 4 で,

$$P(Y=\ell) = {}_4C_\ell \left(\frac{1}{2}\right)^\ell \left(\frac{1}{2}\right)^{4-\ell} \quad (\ell=0, 1, 2, 3, 4)$$

これらから,  $X$  と  $Y$  の確率分布は次のようになる.

$X$	0	1	2	3	計
$P$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	1

$Y$	0	1	2	3	4	計
$P$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$	1

$$\text{よって, } E(X) = 0 \times \frac{1}{8} + 1 \times \frac{3}{8} + 2 \times \frac{3}{8} + 3 \times \frac{1}{8} = \frac{3}{2}$$

$$E(Y) = 0 \times \frac{1}{16} + 1 \times \frac{4}{16} + 2 \times \frac{6}{16} + 3 \times \frac{4}{16} + 4 \times \frac{1}{16} = 2$$

したがって,  $Z$  の平均は,

$$\begin{aligned} E(Z) &= E(100X + 10Y) = 100E(X) + 10E(Y) \\ &= 100 \times \frac{3}{2} + 10 \times 2 = 170 \end{aligned}$$

$$\begin{aligned} E(aX + bY) \\ &= aE(X) + bE(Y) \end{aligned}$$

$$\text{また, } E(X^2) = 0^2 \times \frac{1}{8} + 1^2 \times \frac{3}{8} + 2^2 \times \frac{3}{8} + 3^2 \times \frac{1}{8} = 3$$

$$E(Y^2) = 0^2 \times \frac{1}{16} + 1^2 \times \frac{4}{16} + 2^2 \times \frac{6}{16} + 3^2 \times \frac{4}{16} + 4^2 \times \frac{1}{16} = 5$$

$$\text{より, } X \text{ の分散は, } V(X) = E(X^2) - \{E(X)\}^2 = 3 - \left(\frac{3}{2}\right)^2 = \frac{3}{4}$$

$$Y \text{ の分散は, } V(Y) = E(Y^2) - \{E(Y)\}^2 = 5 - 2^2 = 1$$

$X, Y$  は独立であるから,  $Z$  の分散は,

$$\begin{aligned} V(Z) &= V(100X + 10Y) = 100^2V(X) + 10^2V(Y) \\ &= 100^2 \times \frac{3}{4} + 10^2 \times 1 = 7600 \end{aligned}$$

$$\begin{aligned} X, Y \text{ が独立のとき,} \\ V(aX + bY) \\ &= a^2V(X) + b^2V(Y) \end{aligned}$$

## Focus

一般に, 平均  $E(aX + bY) = aE(X) + bE(Y)$   
 $X$  と  $Y$  が独立ならば, 分散  $V(aX + bY) = a^2V(X) + b^2V(Y)$

## 練習 B2.3

\*\*\*

目が 1, 1, 1, 2, 3, 3 のさいころ A と, 目が 3, 4, 5, 5, 6, 7 のさいころ B を同時に 1 回投げ, さいころ A とさいころ B の出た目の数をそれぞれ縦, 横の長さとする長方形を作る. この長方形の面積  $S$  の平均を求めよ.  $\rightarrow p.B2-13$  [4] [5]



## 例題 B2.4 二項分布の平均と標準偏差

\*\*\*\*

数直線上を動く点 P がある。最初に点 P は原点にあり、さいころを 1 回投げるたびに 3 の倍数の目が出れば正の向きに 2 進み、それ以外の目であれば負の向きに 1 進む。さいころを 72 回投げた結果、3 の倍数の目が出る回数を  $X$ 、点 P の座標を  $Y$  とする。

- (1)  $X$  の平均と標準偏差を求めよ。
- (2)  $Y$  の平均と標準偏差を求めよ。

第2章

**考え方** 1 回の試行で事象  $A$  が起こる確率が  $p$ 、起こらない確率が  $q$  (ただし、 $p+q=1$ ) のとき、 $n$  回の独立試行で事象  $A$  が起こる回数  $X$  は二項分布  $B(n, p)$  に従う。このとき、平均は、 $E(X)=np$ 、標準偏差は、 $\sigma(X)=\sqrt{npq}$  である。

**解答** 1 回投げて、3 の倍数の目が出る確率は、 $\frac{2}{6}=\frac{1}{3}$  ▶ 3 の倍数の目は 3 と 6 の 2 つ  
 それ以外の目が出る確率は、 $1-\frac{1}{3}=\frac{2}{3}$

- (1) 確率変数  $X$  が  $X=k$  となる確率は、

$$P(X=k)={}_{72}C_k\left(\frac{1}{3}\right)^k\left(\frac{2}{3}\right)^{72-k} \quad (k=0, 1, 2, \dots, 72)$$

$X$  は二項分布  $B\left(72, \frac{1}{3}\right)$  に従うから、

$$X \text{ の平均は, } E(X)=72 \times \frac{1}{3}=24$$

$$\text{▶ } E(X)=np$$

$$\text{標準偏差は, } \sigma(X)=\sqrt{72 \times \frac{1}{3} \times \frac{2}{3}}=\sqrt{16}=4$$

$$\text{▶ } \sigma(X)=\sqrt{npq}$$

- (2) 3 の倍数の目が  $X$  回出るとすると、それ以外の目は  $(72-X)$  回出るから、点 P の座標  $Y$  は、

$$Y=2 \times X + (-1) \times (72-X)=3X-72$$

よって、 $Y$  の平均は、

$$\begin{aligned} E(Y) &= E(3X-72) = 3E(X)-72 \\ &= 3 \times 24 - 72 = 0 \end{aligned}$$

▶  $Y$  は  $X$  の 1 次式で表される。

また、 $Y$  の標準偏差は、

$$\sigma(Y)=\sigma(3X-72)=|3|\sigma(X)=3 \times 4=12$$

## Focus

$X$  が二項分布  $B(n, p)$  に従うとき、 $q=1-p$  とすると、  
 平均  $E(X)=np$ 、標準偏差  $\sigma(X)=\sqrt{npq}$

## 練習 B2.4

\*\*

A、B の 2 つの野球チームが 5 試合行う。引き分けはないものとし、各試合で A チームが B チームに勝つ確率は  $\frac{5}{8}$  である。

- (1) A チームが勝った試合数を  $X$  とするとき、 $X$  の平均と標準偏差を求めよ。
- (2) 勝った試合数の 2 乗の勝ち点がもらえるとする。A チームの勝ち点の平均を求めよ。

➡ p.B2-13 [6] [7]

B1

B2

C1

C2

## 解説

Commentary

## 「微分を利用した二項分布の平均と標準偏差」

二項分布  $B(n, p)$  に従う確率変数  $X$  において、平均  $E(X) = np$ 、標準偏差  $\sigma(X) = \sqrt{npq}$  が成り立つ。

ここでは、微分を利用して、これらの公式の成立を証明してみよう。

(証明)  $X$  が二項分布  $B(n, p)$  に従うとき、 $q = 1 - p$  とすると、確率分布は下のようになる。

$X$	0	1	……	$r$	……	$n$	計
$P$	${}_nC_0 p^0 q^n$	${}_nC_1 p^1 q^{n-1}$	……	${}_nC_r p^r q^{n-r}$	……	${}_nC_n p^n q^0$	1

$$\begin{aligned}
 E(X) &= 0 \cdot {}_nC_0 p^0 q^n + 1 \cdot {}_nC_1 p^1 q^{n-1} + \cdots + r \cdot {}_nC_r p^r q^{n-r} + \cdots \\
 &\quad \cdots + n \cdot {}_nC_n p^n q^0 \\
 &= \sum_{k=0}^n k {}_nC_k p^k q^{n-k}
 \end{aligned}$$

$$\text{ここで, } (q + px)^n = \sum_{k=0}^n {}_nC_k q^{n-k} p^k x^k \quad (n \geq 2)$$

を考え、両辺を  $x$  で微分すると、

$$np(q + px)^{n-1} = \sum_{k=1}^n {}_nC_k q^{n-k} p^k k x^{k-1} \quad \cdots \cdots ①$$

もう1度  $x$  で微分すると、

$$n(n-1)p^2(q + px)^{n-2} = \sum_{k=2}^n {}_nC_k q^{n-k} p^k k(k-1) x^{k-2} \quad \cdots \cdots ②$$

①, ②で  $x=1$  とおくと、 $p + q = 1$  より、

$$\text{①から, } np = \sum_{k=1}^n {}_nC_k q^{n-k} p^k k = \sum_{k=0}^n {}_nC_k q^{n-k} p^k k \quad \cdots \cdots ③$$

$$\begin{aligned}
 \text{②から, } n(n-1)p^2 &= \sum_{k=2}^n {}_nC_k q^{n-k} p^k k(k-1) \\
 &= \sum_{k=0}^n {}_nC_k q^{n-k} p^k k(k-1) \quad \cdots \cdots ④
 \end{aligned}$$

$$\begin{aligned}
 \text{よって, ③, ④から, } \sum_{k=0}^n {}_nC_k q^{n-k} p^k k^2 &= n(n-1)p^2 + \sum_{k=0}^n {}_nC_k q^{n-k} p^k k \\
 &= n(n-1)p^2 + np \quad \cdots \cdots ⑤
 \end{aligned}$$

$$\text{③から, } E(X) = np \quad \cdots \cdots ⑥$$

$$\begin{aligned}
 \text{⑤, ⑥から, } V(X) &= E(X^2) - \{E(X)\}^2 = \sum_{k=0}^n k^2 {}_nC_k q^{n-k} p^k - n^2 p^2 \\
 &= n(n-1)p^2 + np - n^2 p^2 \\
 &= np(1-p) = npq \quad \cdots \cdots ⑦
 \end{aligned}$$

$$\text{よって, } \sigma(X) = \sqrt{npq} \quad \cdots \cdots ⑧$$

$$n=1 \text{ のとき, } E(X) = p, \quad V(X) = (1^2 \times p + 0^2 \times q) - p^2 = p(1-p) = pq,$$

$\sigma(X) = \sqrt{pq}$  となり、 $n=1$  のときも⑥, ⑦, ⑧は成り立つ。

以上より、 $E(X) = np$ ,  $V(X) = npq$ ,  $\sigma(X) = \sqrt{npq}$  が成り立つ。

## Think

## 例題

## B2.5 確率変数の平均・標準偏差

\*\*\*\*

袋の中に  $n$  個 ( $n \geq 3$ ) の玉が入っている。そのうちの 2 個は白玉で、残りは黒玉である。この袋から 1 個ずつ玉を取り出していく。ただし、取り出した玉は袋の中に戻さない。白玉がはじめて出るまでに取り出される黒玉の個数  $X$  の平均と標準偏差を求めよ。

## 考え方

たとえば、 $X=3$  となるのは、3 回目まで黒玉が取り出され、4 回目にはじめて白玉が取り出されるときで、その確率は、 $P(X=3) = \frac{n-2}{n} \cdot \frac{n-3}{n-1} \cdot \frac{n-4}{n-2} \cdot \frac{2}{n-3}$  である。

## 解答

最初に袋の中に入っている黒玉の数は  $n-2$  (個) であるから、確率変数  $X$  のとり得る値は、0, 1, 2, 3, …,  $n-2$  である。

また、 $X$  が 0 となる確率は、 $P(X=0) = \frac{2}{n}$  である。

$1 \leq k \leq n-2$  のとき、

$$P(X=k) = \frac{n-2}{n} \cdot \frac{n-3}{n-1} \cdot \frac{n-4}{n-2} \cdots \frac{n-k-1}{n-k+1} \cdot \frac{2}{n-k} = \frac{n-k-1}{n} \cdot \frac{2}{n-1}$$

よって、黒玉の個数  $X$  の平均は、

$$E(X) = 0 \cdot \frac{2}{n} + \sum_{k=1}^{n-2} k \cdot \frac{n-k-1}{n} \cdot \frac{2}{n-1} = \frac{2}{n(n-1)} \sum_{k=1}^{n-2} k(n-k-1)$$

$$= \frac{2}{n(n-1)} \left\{ (n-1) \sum_{k=1}^{n-2} k - \sum_{k=1}^{n-2} k^2 \right\}$$

$$= \frac{2}{n(n-1)} \left\{ (n-1) \cdot \frac{1}{2} (n-2)(n-1) \right.$$

$$\left. - \frac{1}{6} (n-2)(n-1)(2n-3) \right\}$$

$$= \frac{n-2}{n} \left( n-1 - \frac{2n-3}{3} \right) = \frac{n-2}{3}$$

$$\text{また, } E(X^2) = 0^2 \cdot \frac{2}{n} + \sum_{k=1}^{n-2} k^2 \cdot \frac{n-k-1}{n} \cdot \frac{2}{n-1}$$

$$= \frac{2}{n(n-1)} \sum_{k=1}^{n-2} k^2(n-k-1)$$

$$= \frac{2}{n(n-1)} \left\{ (n-1) \sum_{k=1}^{n-2} k^2 - \sum_{k=1}^{n-2} k^3 \right\}$$

$$= \frac{2}{n(n-1)} \left\{ (n-1) \cdot \frac{1}{6} (n-2)(n-1)(2n-3) - \frac{1}{2^2} (n-2)^2 (n-1)^2 \right\}$$

$$= \frac{(n-1)(n-2)}{n} \left( \frac{2n-3}{3} - \frac{n-2}{2} \right) = \frac{(n-1)(n-2)}{6}$$

よって、分散は、

$$V(X) = E(X^2) - \{E(X)\}^2 = \frac{(n-1)(n-2)}{6} - \left( \frac{n-2}{3} \right)^2 = \frac{(n-2)(n+1)}{18}$$

$$\text{したがって、標準偏差は、} \sigma(X) = \sqrt{V(X)} = \frac{\sqrt{2(n-2)(n+1)}}{6}$$

## 練習

## B2.5

\*\*\*

赤い本が 2 冊、青い本が  $n$  冊ある。この  $n+2$  (冊) の本を無作為に 1 冊ずつ選び、本棚に左から並べていく。2 冊の赤い本の間にある青い本の冊数を  $X$  とすると、 $X$  の平均と分散を求めよ。

## 第2章

B1

B2

C1

C2

## Think

## 例題 B2.6 漸化式と平均・分散

\*\*\*\*

- (1) 硬貨を5回投げて、表の出た回数と裏の出た回数の差の絶対値を  $X_0$  とする。確率変数  $X_0$  の平均  $E(X_0)$  と分散  $V(X_0)$  を求めよ。
- (2) (1)の  $X_0$  から始まり、 $4X_n = X_{n-1} + 3$  ( $n=1, 2, \dots$ ) によって定まる確率変数の列  $X_0, X_1, X_2, \dots, X_n, \dots$  がある。 $X_n$  の平均  $E(X_n)$  と分散  $V(X_n)$  を求めよ。

## 考え方

- (1) たとえば、(表, 裏)=(1回, 4回), (4回, 1回) のとき,  $X_0=3$  となる。

また、このときの確率は、 ${}_5C_1\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^4 + {}_5C_4\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^1$  である。

- (2)  $X_n$  は、2項間の漸化式の考え方を利用して求める。

## 解答

- (1) 硬貨を5回投げたとき、表と裏の出る回数、回数の差の絶対値  $X_0$  の値、および、それが起こる確率は次のようになる。

(表, 裏)=(0, 5), (5, 0) のとき,  $X_0=5$  であり、

$$P(X_0=5) = 2 \times {}_5C_0 \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^5 = \frac{1}{2^4}$$

(表, 裏)=(1, 4), (4, 1) のとき,  $X_0=3$  であり、

$$P(X_0=3) = 2 \times {}_5C_1 \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^4 = \frac{5}{2^4}$$

(表, 裏)=(2, 3), (3, 2) のとき,  $X_0=1$  であり、

$$P(X_0=1) = 2 \times {}_5C_2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^3 = \frac{10}{2^4}$$

$$\begin{aligned} & {}_5C_5 \left(\frac{1}{2}\right)^5 \left(\frac{1}{2}\right)^0 \\ &= {}_5C_0 \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^5 \\ & \text{(表, 裏)=(4, 1), (3, 2)} \\ & \text{のときも同様} \end{aligned}$$

よって、平均は、 $E(X_0) = 5 \times \frac{1}{2^4} + 3 \times \frac{5}{2^4} + 1 \times \frac{10}{2^4} = \frac{15}{8}$

また、 $E(X_0^2) = 5^2 \times \frac{1}{2^4} + 3^2 \times \frac{5}{2^4} + 1^2 \times \frac{10}{2^4} = 5$  より、分散は、

$$V(X_0) = E(X_0^2) - \{E(X_0)\}^2 = 5 - \left(\frac{15}{8}\right)^2 = \frac{95}{64}$$

- (2)  $4X_n = X_{n-1} + 3$  は、 $X_n - 1 = \frac{1}{4}(X_{n-1} - 1)$  と変形できる。

よって、 $X_n - 1 = \left(\frac{1}{4}\right)^n (X_0 - 1)$  より、 $X_n = \left(\frac{1}{4}\right)^n X_0 - \left(\frac{1}{4}\right)^n + 1$

したがって、

$$\begin{aligned} \text{平均は、} E(X_n) &= \left(\frac{1}{4}\right)^n E(X_0) - \left(\frac{1}{4}\right)^n + 1 = \left(\frac{1}{4}\right)^n \cdot \frac{15}{8} - \left(\frac{1}{4}\right)^n + 1 \\ &= \frac{7}{8} \left(\frac{1}{4}\right)^n + 1 = \frac{7}{2^{2n+3}} + 1 \end{aligned}$$

$$\text{分散は、} V(X_n) = \left\{ \left(\frac{1}{4}\right)^n \right\}^2 V(X_0) = \left\{ \left(\frac{1}{4}\right)^n \right\}^2 \cdot \frac{95}{64} = \frac{95}{2^{4n+6}}$$

特性方程式  $4\alpha = \alpha + 3$   
より、 $\alpha = 1$

## 練習 B2.6

\*\*\*\*

赤玉が3個、白玉が2個、青玉が1個入っている袋がある。この袋から3個の玉を同時に取り出すとき、取り出された玉の色が何種類であるかを確率変数  $X_0$  で表す。 $X_0$  から始まり、 $X_n = 3X_{n-1} + 2$  ( $n=1, 2, \dots$ ) によって定まる確率変数の列  $X_0, X_1, X_2, \dots, X_n, \dots$  について、 $X_n$  の平均  $E(X_n)$  と分散  $V(X_n)$  を求めよ。

## Step Up

確率分布 ▶▶ 解答編 p.B2-9

## 第2章

\*\*

1

p.B2-7

0, 1, 2 のいずれかの値をとる確率変数  $X$  の期待値が 1, 分散が  $\frac{1}{2}$  であるとする. この確率変数  $X$  の確率分布を求めよ. (宮崎大)

\*\*\*

2

p.B2-7

白玉 4 個と赤玉 2 個が入った箱がある. 1 個のさいころを投げて, 1 か 2 の目が出たら, この箱から玉を 4 個取り出し, 1, 2 以外の目が出たら, この箱から玉を 2 個取り出す. 取り出した玉のうちの白玉の個数を  $X$  とする. 確率変数  $X$  の平均  $E(X)$ , 標準偏差  $\sigma(X)$  を求めよ.

\*\*\*

3

p.B2-7

確率変数  $X$  は  $n$  個の値  $1, 3, 5, \dots, 2n-1$  をとるものとする.  $X$  がそれぞれの値を等しい確率でとるとき,  $2X+3$  の平均と分散を求めよ. (茨城大)

\*\*

4

p.B2-8

袋の中に [1], [3], [5], [7] のカードがそれぞれ 4 枚, 3 枚, 2 枚, 1 枚ずつ入っている. この袋から 1 枚取り出しては袋に戻す試行を 5 回繰り返し,  $k$  回目 ( $k=1, 2, \dots, 5$ ) に出たカードの番号が  $X_k$  ならば  $kX_k$  が得点となる. 得点の合計の平均と分散を求めよ.

\*\*\*

5

p.B2-7

p.B2-8

赤玉 5 個と白玉 3 個が入っている袋から, 4 個の玉を同時に取り出すとき, 赤玉の個数を  $X$ , 白玉の個数を  $Y$  とする.

- (1)  $Z=3X+Y$  とするとき,  $Z$  の平均  $E(Z)$ , 標準偏差  $\sigma(Z)$  を求めよ.
- (2)  $T=XY$  とするとき,  $T$  の平均  $E(T)$  を求めよ.

\*\*\*

6

p.B2-9

確率変数  $X$  は二項分布  $B(n, p)$  に従い, その平均  $m$  は  $m=2$  で, 標準偏差  $\sigma$  は  $\sigma=\frac{\sqrt{6}}{2}$  である.

- (1)  $n$  と  $p$  の値を求めよ.
- (2)  $|X-m|>\sigma$  となる確率を求めよ.
- (3)  $X^2+2X-3$  の平均  $E(X^2+2X-3)$  を求めよ.

\*\*

7

p.B2-9

袋の中に 100 個の玉が入っており, そのうち  $2a$  個が赤玉である. この袋から無作為に 1 個の玉を取り出し, 色を調べてからもとに戻す. この操作を  $n$  回繰り返すとすると, 赤玉を取り出した回数  $X$  の平均は  $\frac{16}{5}$ , 分散は  $\frac{64}{25}$  であるという. このとき,  $a$  と  $n$  の値を求めよ.

## Story ストーリー

## 「偏差値について」

みなさんは「偏差値」という言葉を聞いたことがあるかと思う。ここでは、偏差値の定義とその結果から何が言えるのかを考えてみよう。

偏差値は、満点や平均点が異なるテストにおいて成績の良し悪しを統一的に評価する客観的な尺度として用いられる。その定義は、

「受験者の得点の平均が 50 点、標準偏差が 10 点になるように全体のデータを変換した後の得点」である。

具体例を用いて考えてみよう。ある数学のテストの得点の平均は 57 点、標準偏差は 18.2 点とする。そこで、

- (i) 全員の点数から 57 点を引く(平均が 0 になり、標準偏差はそのまま)
  - (ii) その結果を  $\frac{10}{18.2}$  倍する(平均は  $0 \cdot \frac{10}{18.2} = 0$  のまま、標準偏差が 10 になる)
  - (iii) その結果に 50 を加える(平均が 50 になり、標準偏差はそのまま)
- とすると、変換したデータの平均は 50、標準偏差は 10 になる。得点を  $x$ 、変換後のデータを  $z$  として、この変換を式でまとめると、

$$z = \frac{10}{18.2}(x - 57) + 50$$

となるので、

$$\text{得点が 100 点} \quad \cdots \quad \text{偏差値は,} \quad \frac{10}{18.2}(100 - 57) + 50 \div 73.6$$

$$\text{得点が 60 点} \quad \cdots \quad \text{偏差値は,} \quad \frac{10}{18.2}(60 - 57) + 50 \div 51.6$$

$$\text{得点が 10 点} \quad \cdots \quad \text{偏差値は,} \quad \frac{10}{18.2}(10 - 57) + 50 \div 24.2$$

となる。

変換式をまとめて整理すれば、平均  $m$ 、標準偏差  $s$  ( $s \neq 0$ ) の試験で  $X$  点を取った場合の偏差値  $Z$  は、

$$Z = 10 \cdot \frac{X - m}{s} + 50$$

である。自分の得点が平均点と同じであれば偏差値は 50 である。偏差値 60 とは、「平均 + 標準偏差」の点を取ったということである。先程の例では、 $57 + 18.2 = 75.2$  (点) 取ると偏差値が 60 になる。逆に言えば、得点、偏差値、平均点がわかれば、標準偏差と分散を逆算することもできる。

さて、ここで次の問題を考えてみよう。

- (1) 偏差値が 100 を超えることはあるか。
- (2) 偏差値が 0 より小さくなることはあるか。

まずは、(1)について考察してみよう。

$x$  人 ( $x \geq 2$ ) が試験を受けて、そのうち 1 人が 100 点、残りの  $x-1$  人が 0 点であるとする、平均は  $\frac{100}{x}$  点で、分散は、

$$\frac{1}{x} \left\{ \left( 100 - \frac{100}{x} \right)^2 + (x-1) \left( \frac{100}{x} \right)^2 \right\} = 100^2 \cdot \frac{1}{x} \left( 1 - \frac{1}{x} \right)$$

であるから、次の式が成り立つとする。

$$10 \cdot \frac{100 - \frac{100}{x}}{100 \sqrt{\frac{1}{x} \left( 1 - \frac{1}{x} \right)}} + 50 \geq 100$$

これを解くと、 $x \geq 2$  より、 $x \geq 26$

つまり、26 人以上で 100 点が 1 人、残りが 0 点であれば、100 点を取った人は偏差値が 100 を超えることになる。

(練習問題)

(2)について、 $x$  人 ( $x \geq 2$ ) が試験を受けて、そのうち 1 人が 0 点、残りの  $x-1$  人が 100 点であるとする、26 人以上で 0 点が 1 人、残りが 100 点であれば、0 点を取った人は偏差値が 0 以下になる。このことを確かめよ。

さて、偏差値から何がわかるだろうか？仮に、多くの受験者がいる試験でその得点分布が正規分布に非常に近いとする。平均 0、標準偏差 1 に標準化した正規分布での  $X$  という値は、平均 50、標準偏差 10 に直せば、 $Z = 50 + 10X$  になる。

$X = \frac{1}{10}(Z - 50)$  であるから、正規分布表により偏差値 60 以上の割合は、

$P(X \geq 1) = 0.5 - 0.3413 = 0.1587$  で全体の約 16% になる。また、偏差値 70 以上の割合は、 $P(X \geq 2) = 0.5 - 0.4772 = 0.0228$  で全体の 2.28% となる。このことは、全体の得点分布が正規分布に近いことを仮定すれば、偏差値から全体でのおおよその位置がつかめることになる。

しかしながら、実際は得点分布がきれいな正規分布に近くなるとは限らない。左右非対称型やフタコブラクダ型など、偏差値 50 の人でも全体の真ん中にいるとは限らない。

とは言え、偏差値というのは明確でわかりやすいものであるから、たとえば、理科や社会の試験などの選択科目の試験では、偏差値の大小で順位づけをすることは客観的な基準となる。

# まとめ 2 正規分布

## 1. 連続的な確率変数

連続的な値をとる確率変数  $X$  を連続型確率変数といい、これまでに扱った飛び飛びの値をとる確率変数  $X$  を離散型確率変数という。

連続型確率変数  $X$  が  $a$  以上  $b$  以下の値をとる確率を  $P(a \leq X \leq b)$  で表す。

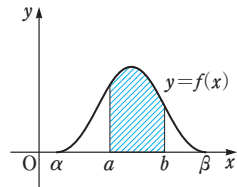
### 確率密度関数と分布曲線

$\alpha \leq X \leq \beta$  の範囲にある連続的な確率変数  $X$  について、次の性質の関数  $f(x)$  を対応させる。

[1]  $f(x) \geq 0$

[2] 確率  $P(a \leq X \leq b)$  は図の斜線部分の面積に等しい。

すなわち、
$$P(a \leq X \leq b) = \int_a^b f(x) dx$$



この関数  $f(x)$  を  $X$  の確率密度関数、曲線  $y=f(x)$  を分布曲線という。  
連続的な確率変数  $X$  のとり得る値の範囲が  $\alpha \leq X \leq \beta$  のとき、

$$P(\alpha \leq X \leq \beta) = \int_{\alpha}^{\beta} f(x) dx = 1$$

((曲線  $y=f(x)$ ,  $x$  軸, 直線  $x=\alpha$ ,  $x=\beta$  で囲まれる部分全体の面積) = 1)

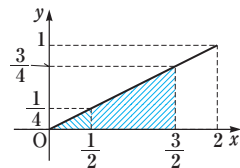
$$P(X=a)=0$$

(例) 確率変数  $X$  の確率密度関数が,  $f(x) = \frac{1}{2}x$  ( $0 \leq x \leq 2$ ) であるとき、

$$P\left(0 \leq X \leq \frac{1}{2}\right) = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{4} = \frac{1}{16}$$

$$P\left(\frac{1}{2} \leq X < \frac{3}{2}\right) = P\left(0 \leq X < \frac{3}{2}\right) - P\left(0 \leq X \leq \frac{1}{2}\right)$$

$$= \frac{1}{2} \times \frac{3}{2} \times \frac{3}{4} - \frac{1}{16} = \frac{1}{2}$$



【注】  $P(X=a)=0$  より,  $P\left(X=\frac{3}{2}\right)=0$  であるから,  $P\left(\frac{1}{2} \leq X < \frac{3}{2}\right) = P\left(\frac{1}{2} \leq X \leq \frac{3}{2}\right)$

### 確率密度関数と平均・分散・標準偏差

$a \leq X \leq b$  の範囲にある連続的な確率変数  $X$  について、確率密度関数が  $f(x)$  であるとき、平均  $m=E(X)$ 、分散  $V(X)$ 、標準偏差  $\sigma(X)$  は、

$$m=E(X) = \int_a^b x f(x) dx$$

$$V(X) = \int_a^b (x-m)^2 f(x) dx = \int_a^b x^2 f(x) dx - m^2$$

$$\sigma(X) = \sqrt{V(X)}$$



## 2. 正規分布

確率変数  $X$  のとり得る値が実数全体で、 $X$  の確率密度関数が、

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} \quad (e \text{ は無理数で、} e=2.71828\cdots)$$

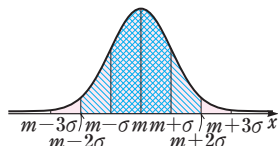
であるとき、この  $X$  の確率分布を、平均  $m$ 、標準偏差  $\sigma$  の正規分布といい、 $N(m, \sigma^2)$  で表す。このとき、確率変数  $X$  は正規分布  $N(m, \sigma^2)$  に従うという。

正規分布  $N(m, \sigma^2)$  における  $y=f(x)$  のグラフを正規分布曲線といい、これは  $x=m$  に関して対称で、次の性質をもつ確率分布である。

$$P(m-\sigma \leq X \leq m+\sigma) \doteq 0.6827$$

$$P(m-2\sigma \leq X \leq m+2\sigma) \doteq 0.9545$$

$$P(m-3\sigma \leq X \leq m+3\sigma) \doteq 0.9973$$



第2章

### 標準正規分布

確率変数  $X$  が正規分布  $N(m, \sigma^2)$  に従うとき、 $Z = \frac{X-m}{\sigma}$  とおくと、 $Z$  は

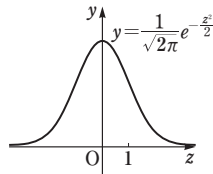
$X$  の1次式であるから、確率変数  $Z$  の平均  $E(Z)$ 、標準偏差  $\sigma(Z)$  は、

$$E(Z) = \frac{1}{\sigma} \{E(X) - m\} = \frac{1}{\sigma} (m - m) = 0$$

$$\sigma(Z) = \frac{1}{|\sigma|} \cdot \sigma(X) = \frac{1}{\sigma} \cdot \sigma = 1$$

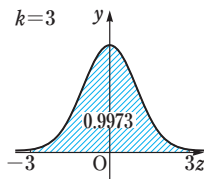
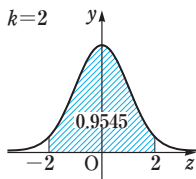
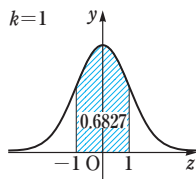
より、 $Z$  の確率分布は、平均 0、標準偏差 1 の正規分布  $N(0, 1)$  となる。正規分布  $N(0, 1)$  を標準正規分布と

いい、その確率密度関数は、 $f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$



標準正規分布は、次の性質をもつ確率分布である。

- ①  $y=f(z)$  のグラフは、 $y$  軸に関して対称な山型の曲線である。
- ②  $Z$  の値が  $-k \leq Z \leq k$  の範囲にある確率は、次の図のような斜線部分の面積となり、 $k$  が大きくなると 1 に近づく。



【注】 標準正規分布  $N(0, 1)$  に従う確率変数  $Z$  について、 $\alpha > 0$ ,  $\beta > 0$  とすると、

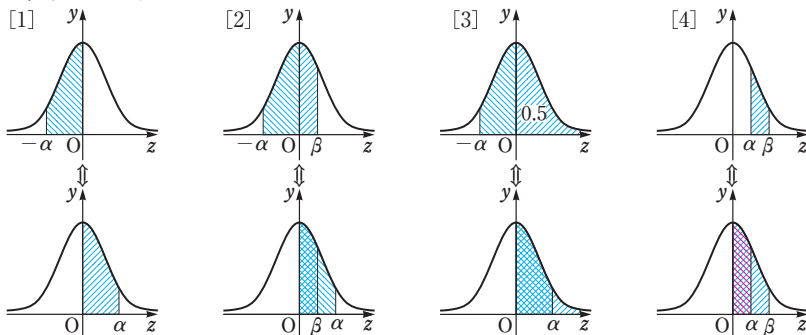
$$[1] \quad P(-\alpha \leq Z \leq 0) = P(0 \leq Z \leq \alpha)$$

$$[2] \quad P(-\alpha \leq Z \leq \beta) = P(0 \leq Z \leq \alpha) + P(0 \leq Z \leq \beta)$$

$$[3] \quad P(-\alpha \leq Z) = P(0 \leq Z \leq \alpha) + 0.5$$

$$[4] \quad P(\alpha \leq Z \leq \beta) = P(0 \leq Z \leq \beta) - P(0 \leq Z \leq \alpha) \quad (\alpha < \beta \text{ のとき})$$

標準正規分布  $N(0, 1)$  において、 $P(0 \leq Z \leq u)$  ( $u$  は正の実数) のことを、 $p(u)$  のように表すこともある。



(例) 確率変数  $Z$  が標準正規分布  $N(0, 1)$  に従うとき、確率  $P(Z > 2)$  の値は、

$$P(Z > 2) = 0.5 - P(0 \leq Z \leq 2) = 0.5 - \frac{0.9545}{2} = 0.02275$$

標準正規分布  $N(0, 1)$  に従う確率変数  $Z$  に対して、確率  $P(0 \leq Z \leq u)$  を、 $u$  のいろいろな値に対して計算して表にまとめたものを正規分布表という。

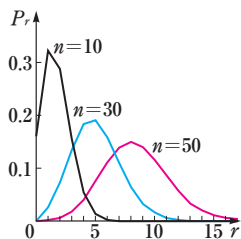
### 3. 正規分布による二項分布の近似

二項分布  $B(n, p)$  に従う確率変数  $X$  について、 $X=r$  となる確率を  $P_r$  とする。

$P_r = {}_n C_r p^r q^{n-r}$  ( $q=1-p$ ) を、 $p=\frac{1}{6}$  のとき、 $n=10, 30, 50$  の各場合について

計算し、点  $(r, P_r)$  を定めて折れ線グラフで表すと、右の図のようになり、 $n$  が大きくなると、二項分布のグラフはほぼ左右対称で、正規分布曲線に近くなる。

一般に、確率変数  $X$  が二項分布  $B(n, p)$  に従うとき、 $n$  が大きくなるとグラフは正規分布曲線に近くなり、 $X$  はほぼ正規分布に従うことが知られている。したがって、 $X$  の



平均は  $m = E(X) = np$ 、標準偏差は  $\sigma = \sigma(X) = \sqrt{npq}$  ( $q=1-p$ )

より、 $Z = \frac{X-m}{\sigma} = \frac{X-np}{\sqrt{npq}}$  とおくと、 $Z$  はほぼ標準正規分布に従う。

以上より、確率変数  $X$  が二項分布  $B(n, p)$  に従うとき、 $n$  が大きければ、

$Z = \frac{X-np}{\sqrt{npq}}$  ( $q=1-p$ ) は、ほぼ標準正規分布  $N(0, 1)$  に従う。

## Check!

## 第2章

\*

7

確率変数  $X$  が区間  $0 \leq X \leq 4$  の任意の値をとり、その確率密度関数が  $f(x) = k(5-x)$  であるとき、 $k$  の値を求めよ。また、 $P(1 \leq X \leq 3)$  を求めよ。

\*

8

確率変数  $X$  の確率密度関数が  $f(x) = \frac{1}{2}x$  ( $0 \leq x \leq 2$ ) のとき、確率変数  $X$  の平均、分散、標準偏差を求めよ。

\*

9

確率変数  $X$  が正規分布  $N(24, 8^2)$  に従うとき、 $P(X \leq 16)$ 、および  $P(8 \leq X \leq 32)$  を求めよ。  
ただし、 $P(|X - 24| \leq 8) = 0.6827$ 、 $P(|X - 24| \leq 16) = 0.9545$  として計算せよ。

\*\*

10

正規分布  $N(18, 6^2)$  に従う確率変数  $X$  について、次の等式が成り立つ。  
このとき、正規分布表を用いて  $a$ 、 $b$  の値を求めよ。  
(1)  $P(X \leq a) = 0.9032$   
(2)  $P(12 \leq X \leq b) = 0.7745$

\*\*

11

ある学校で 500 人の生徒にテストを行ったところ、その得点  $X$  は平均 48 点、標準偏差 16 点の正規分布に従った。正規分布表を用いて次の問いに答えよ。ただし、得点は整数とする。  
(1) 76 点以上の生徒は約何人いるか。  
(2) 得点が高い方から 50 人の中に入るには、約何点以上であればよいか。

\*\*

12

さいころを 720 回振るとき、1 の目が出る回数を  $X$  とする。 $X$  が 100 以上 130 以下の範囲にある確率  $P(100 \leq X \leq 130)$  を、正規分布表を用いて求めよ。

▶▶ 解答編 p.B2-15

7  $k = \frac{1}{12}$ ,  $P(1 \leq X \leq 3) = \frac{1}{2}$       8 平均  $\frac{4}{3}$ , 分散  $\frac{2}{9}$ , 標準偏差  $\frac{\sqrt{2}}{3}$

9  $P(X \leq 16) = 0.15865$ ,  $P(8 \leq X \leq 32) = 0.8186$       10 (1)  $a = 25.8$  (2)  $b = 27$

11 (1) 約 20 人 (2) 約 69 点以上      12 0.8185

B1

B2

C1

C2

## 例題 B2.7 確率密度関数と平均・標準偏差

\*\*\*\*

確率変数  $X$  が区間  $0 \leq X \leq 10$  の任意の値をとり、その確率密度関数が  $f(x) = kx(10-x)$  ( $k$  は正の定数) である。

- (1)  $k$  の値を求めよ。 (2) 確率  $P(3 \leq X \leq 6)$  を求めよ。  
 (3) 平均  $m$  と標準偏差  $\sigma$  を求めよ。

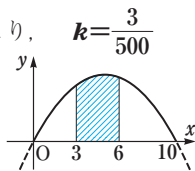
- 考え方** (1) 確率密度関数  $y=f(x)$  ( $\geq 0$ ) のグラフと  $x$  軸で囲まれる部分の面積は 1 である。  
 (2) 確率  $P(a \leq X \leq b) = \int_a^b f(x) dx$  である。  
 (3) 平均  $m = \int_0^{10} xf(x) dx$ , 分散  $V(X) = \int_0^{10} x^2 f(x) dx - m^2$  (詳しくは **注** 参照)

- 解答** (1) 確率変数  $X$  がとり得る値の範囲は  $0 \leq X \leq 10$  で、  
 $k > 0$  より、

$$P(0 \leq X \leq 10) = \int_0^{10} f(x) dx = \int_0^{10} kx(10-x) dx = 1$$

よって、 $\int_0^{10} kx(10-x) dx = k \left[ 5x^2 - \frac{1}{3}x^3 \right]_0^{10} = \frac{500}{3}k = 1$  より、 $k = \frac{3}{500}$

$$(2) P(3 \leq X \leq 6) = \int_3^6 \frac{3}{500}x(10-x) dx = \frac{3}{500} \left[ 5x^2 - \frac{1}{3}x^3 \right]_3^6 = \frac{3}{500} \{ (180-72) - (45-9) \} = \frac{54}{125}$$



$$(3) m = \int_0^{10} xf(x) dx = \int_0^{10} \frac{3}{500}x^2(10-x) dx = \frac{3}{500} \left[ \frac{10}{3}x^3 - \frac{1}{4}x^4 \right]_0^{10} = 5$$

$$\begin{aligned} \text{分散は、} V(X) &= \int_0^{10} (x-m)^2 f(x) dx = \int_0^{10} x^2 f(x) dx - m^2 \\ &= \int_0^{10} \frac{3}{500}x^3(10-x) dx - m^2 = \frac{3}{500} \left[ \frac{10}{4}x^4 - \frac{1}{5}x^5 \right]_0^{10} - m^2 \\ &= \frac{3}{500} \left( \frac{10^5}{4} - \frac{10^5}{5} \right) - 5^2 = \frac{3}{500} \cdot \frac{10^5}{20} - 25 = 30 - 25 = 5 \end{aligned}$$

したがって、 $\sigma = \sqrt{V(X)} = \sqrt{5}$

## Focus

$X$  の確率密度関数が  $f(x)$  のとき、

$X$  が  $a \leq X \leq b$  にある確率は、 $P(a \leq X \leq b) = \int_a^b f(x) dx$

**注**  $\alpha \leq X \leq \beta$  にある確率変数  $X$  の平均は  $m = \int_\alpha^\beta xf(x) dx$  で、 $\int_\alpha^\beta f(x) dx = 1$  である。

$$\text{分散 } V(X) = \int_\alpha^\beta (x-m)^2 f(x) dx = \int_\alpha^\beta x^2 f(x) dx - 2m \int_\alpha^\beta xf(x) dx + m^2 \int_\alpha^\beta f(x) dx$$

ここに、 $\int_\alpha^\beta xf(x) dx = m$ ,  $\int_\alpha^\beta f(x) dx = 1$  を代入して、 $V(X) = \int_\alpha^\beta x^2 f(x) dx - m^2$

## 練習

## B2.7

確率変数  $X$  が区間  $1 \leq X \leq 9$  の任意の値をとり、その確率密度関数が

$f(x) = k(x+3-2|x-3|)$  ( $k$  は正の定数) である。

\*\*

- (1)  $k$  の値を求めよ。 (2) 確率  $P(2 \leq X \leq 6)$  を求めよ。  
 (3) 平均  $m$  と標準偏差  $\sigma$  を求めよ。

→ p.B2-25 8

## 例題 B2.8 正規分布の標準化(1)

\*\*\*\*

大勢の受験生が受けた2つの試験の平均点はそれぞれ55.8, 78.2, 標準偏差はそれぞれ10.2, 6.4であった。Aは前者の試験を受けて72点, Bは後者の試験を受けて86点であり, どちらの試験の得点も正規分布に従うとき, AとBのどちらが, より学力が優れていると考えられるか。

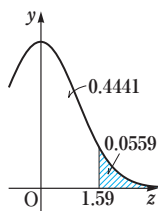
## 第2章

**考え方** 確率変数  $X$  が正規分布  $N(m, \sigma^2)$  に従うとき,  $Z = \frac{X-m}{\sigma}$  とおくと,  $Z$  は標準正規分布  $N(0, 1)$  に従う。A, B の得点を超える受験生の割合を正規分布表を用いて調べると, A, B の学力の位置付けが把握できる。

**解答** 前者の試験の得点を  $X$  とすると,  $X$  は正規分布  $N(55.8, 10.2^2)$  に従うから,

$$Z_1 = \frac{X-55.8}{10.2} \text{ とおくと, } Z_1 \text{ は標準正規分布 } N(0, 1) \text{ に従う。}$$

$$\begin{aligned} \text{よって, } P(X \geq 72) &= P\left(Z_1 \geq \frac{72-55.8}{10.2}\right) \\ &\doteq P(Z_1 \geq 1.59) \\ &= 0.5 - 0.4441 \\ &= 0.0559 \end{aligned}$$



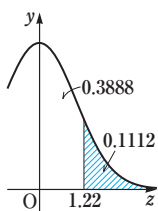
$$\frac{72-55.8}{10.2} = \frac{162}{102} = 1.588\ldots$$

$$\begin{aligned} P(Z_1 \geq 1.59) \\ &= 0.5 - P(0 \leq Z_1 \leq 1.59) \end{aligned}$$

後者の試験の得点を  $Y$  とすると,  $Y$  は正規分布  $N(78.2, 6.4^2)$  に従うから,

$$Z_2 = \frac{Y-78.2}{6.4} \text{ とおくと, } Z_2 \text{ は標準正規分布 } N(0, 1) \text{ に従う。}$$

$$\begin{aligned} \text{よって, } P(Y \geq 86) &= P\left(Z_2 \geq \frac{86-78.2}{6.4}\right) \\ &\doteq P(Z_2 \geq 1.22) \\ &= 0.5 - 0.3888 \\ &= 0.1112 \end{aligned}$$



$$\frac{86-78.2}{6.4} = \frac{78}{64} = 1.218\ldots$$

$$\begin{aligned} P(Z_2 \geq 1.22) \\ &= 0.5 - P(0 \leq Z_2 \leq 1.22) \end{aligned}$$

したがって,  $0.0559 < 0.1112$  から, A の方が試験を受けた各集団の中で学力が上位にあると考えられる。

よって, A の方が学力が優れていると判断できる。

A は上位約 5.59%,  
B は上位約 11.12%  
の生徒であると考えられる。

## Focus

確率変数  $X$  が正規分布  $N(m, \sigma^2)$  に従うとき,  $Z = \frac{X-m}{\sigma}$  で定まる確率変数  $Z$  は標準正規分布  $N(0, 1)$  に従う

## 練習 B2.8

\*\*

800 人の受験生が受けた英語, 国語, 数学の試験の得点は正規分布に従い, 平均点は, それぞれ 54.8, 60.4, 48.3 で, 標準偏差は, それぞれ 12.4, 11.2, 16.1 であった。A の得点が英語 72 点, 国語 78 点, 数学 68 点であるとき, どの教科の成績順位が最も高いといえるか。

→ p.B2-25 ⑨

B1  
B2  
C1  
C2

## 例題 B2.9 正規分布の標準化(2)

\*\*\*\*

得点(整数)が正規分布に従う, ある試験において, 平均  $m=56$ , 標準偏差  $\sigma=12$  で, 42 点以上 78 点以下の人数が 520 人であった.

- (1) 受験生の総数はおよそ何人か.
- (2) 得点  $X$  が  $X \geq m + 1.5\sigma$  を満たすとき, 評価を優とする. 優はおおよそ何人か.
- (3) 成績順位が上から 70 位の A 君の得点  $T$  より 10 点低い B 君の成績順位はおおよそ何位か.

## 考え方

- (1) 正規分布  $N(56, 12^2)$  を標準正規分布  $N(0, 1)$  に変換して考える.
- (3) 成績順位 70 位以上の受験生の割合は,  $70 \div (\text{受験者総数})$  である.

## 解答

- (1) 得点  $X$  は正規分布  $N(56, 12^2)$  に従うから,  $Z = \frac{X-56}{12}$  とおくと,  $Z$  は標準正規分布  $N(0, 1)$  に従う.

$$\begin{aligned} \text{よって, } P(42 \leq X \leq 78) &= P\left(\frac{42-56}{12} \leq Z \leq \frac{78-56}{12}\right) \\ &\doteq P(-1.17 \leq Z \leq 1.83) \\ &= P(0 \leq Z \leq 1.17) + P(0 \leq Z \leq 1.83) \\ &= 0.3790 + 0.4664 = 0.8454 \end{aligned}$$

受験生の総数を  $n$  とすると,  $n \times 0.8454 = 520$  より,  $n = \frac{520}{0.8454} = 615.09 \cdots$  から, 受験生の総数は約 615 人

全体のおおよそ 84.54% が 42 点以上 78 点以下である.

- (2)  $X \geq m + 1.5\sigma$  より,  $Z = \frac{X-m}{\sigma} \geq 1.5$  であるから,

$$P(X \geq m + 1.5\sigma) = P(Z \geq 1.5) = 0.5 - 0.4332 = 0.0668$$

よって, 優の評価だったのは,  $615 \times 0.0668 = 41.082$  より, 約 41 人

- (3) 70 位以上の受験生の割合は  $\frac{70}{615} \doteq 0.1138$  であるから,

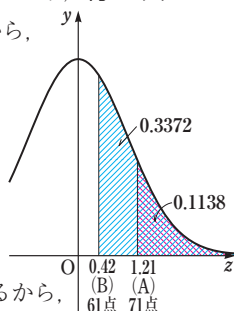
$$\begin{aligned} P\left(0 \leq Z \leq \frac{T-56}{12}\right) &= 0.5 - 0.1138 = 0.3862 \text{ より,} \\ \frac{T-56}{12} \doteq 1.21 \text{ から, } T &\doteq 56 + 12 \times 1.21 = 70.52 \end{aligned}$$

つまり, A 君の得点  $T$  は約 71 点である.

したがって, B 君の得点は,  $71 - 10 = 61$  (点)

$$P\left(Z \geq \frac{61-56}{12}\right) \doteq P(Z \geq 0.42) = 0.5 - 0.1628 = 0.3372$$

より, B 君の順位は上から 0.3372 の割合の位置にあるから,  $615 \times 0.3372 = 207.378$  より, B 君は上から約 207 位



## 練習

## B2.9

\*\*\*

ある入学試験で 300 人の募集定員(合格者定数)に対して 1500 人が受験した. 試験は 500 点満点で, 受験生の得点は整数であり, 平均 232 点, 標準偏差 54 点の正規分布に従った.

- (1) 合格最低点は約何点と考えられるか.
- (2) 合格最低点とそれより 5 点低い得点の間には, 受験生が約何人いると考えられるか.

→ p.B2-25 ⑩

## 例題 B2.10 二項分布と正規分布(1)

\*\*\*\*

ある植物の種の発芽率は 60% である. この種を 600 個まくとする.

- (1) 発芽した種の数  $X$  が 340 以上となる確率を求めよ.
- (2) 発芽した種の数  $Y$  が  $Y \geq \alpha$  の範囲にある確率が 0.7 以上となるような整数  $\alpha$  の最大値を求めよ.

## 第2章

**考え方** 600 個の種をまき, 1 個の種が発芽する確率は,  $\frac{60}{100} = \frac{3}{5}$  であるから,  $X$  は二項分布  $B(600, \frac{3}{5})$  に従う.

- (1) 標準正規分布曲線は直線  $x=0$  に関して対称なグラフであるから, たとえば, 確率  $P(Z \geq -1.2)$  の値は,  $P(0 \leq Z \leq 1.2) + 0.5$  で求める.
- (2)  $P(Z \geq \frac{\alpha-360}{12}) \geq 0.7 = 0.5 + 0.2$  より,  $\alpha - 360 < 0$  で,  $P(0 \leq Z \leq -\frac{\alpha-360}{12}) \geq 0.2$  となる  $\alpha$  の最大値を求める.

**解答** 600 個の種をまき, 発芽率は  $\frac{3}{5}$  であるから,  $X$  は二項分布  $B(600, \frac{3}{5})$  に従う.

よって,  $Z = \frac{X - 600 \times \frac{3}{5}}{\sqrt{600 \times \frac{3}{5} \times (1 - \frac{3}{5})}} = \frac{X - 360}{12}$  とおくと,  $Z$  の分布は標準正規分布  $N(0, 1)$  とみなせる.

$$(1) P(X \geq 340) = P\left(Z \geq \frac{340 - 360}{12}\right) = P(Z \geq -1.67) = 0.4525 + 0.5 = 0.9525$$

したがって, 求める確率は,

$$(2) P(Y \geq \alpha) = P\left(Z \geq \frac{\alpha - 360}{12}\right) \geq 0.7 = 0.5 + 0.2$$

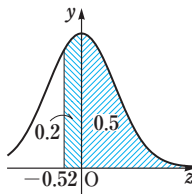
$$P\left(Z \geq \frac{\alpha - 360}{12}\right) > 0.5 \text{ より,}$$

$$P\left(0 \leq Z \leq -\frac{\alpha - 360}{12}\right) \geq 0.2$$

$$\text{であるから, } -\frac{\alpha - 360}{12} > 0.52 \text{ より, } \alpha < 353.76$$

したがって,  $\alpha$  の最大値は, 353

$X$  が二項分布  $B(n, p)$  に従うとき,  $n$  が大きければ,  
 $Z = \frac{X - np}{\sqrt{npq}}$  ( $q = 1 - p$ )  
 は, ほぼ標準正規分布  $N(0, 1)$  に従う.



$$Z = \frac{Y - 360}{12} \geq \frac{\alpha - 360}{12}$$

$$P(0 \leq Z \leq 0.52) = 0.1985$$

$$P(0 \leq Z \leq 0.53) = 0.2019$$

## Focus

二項分布  $B(n, p)$  に従う確率変数  $X$  の  
 平均  $m = np$ , 標準偏差  $\sigma = \sqrt{np(1-p)}$

練習  
B2.10

\*\*

1 問あたりの正答率が 0.8 である問題を 400 問解答し, その正答数を  $X$  とする.  
 $X \leq \alpha$  の範囲にある確率が 0.4 以下となるような整数  $\alpha$  の最大値を求めよ.

→ p.B2-25 (11) (12)

B1  
B2  
C1  
C2

## 例題 B2.11 二項分布と正規分布(2)

\*\*\*\*

座標平面上で点 A が原点 O を出発し、さいころを投げて 1, 2, 3, 4 の目が出れば右に 1 進み, 5, 6 の目が出れば上に 1 進むとする.

- (1) さいころを何回か投げ、点 A が点 (4, 3) に到達する確率を求めよ.  
 (2) さいころを 200 回投げたとき、点 A が  $\{(x, y) | 125 \leq x \leq 140, 0 \leq y \leq 200\}$  の範囲にある確率を求めよ.

## 考え方

- (1) 点  $(a, b)$  に到達する確率は、 ${}_{a+b}C_a \left(\frac{4}{6}\right)^a \left(\frac{2}{6}\right)^b$  である.  
 (2) 確率変数  $y$  の範囲は  $0 \leq y \leq 200$  より、確率  $P(0 \leq y \leq 200) = 1$  である.  
 したがって、 $x$  座標が  $125 \leq x \leq 140$  である確率を求める.

## 解答

- (1) 点 A が点 (4, 3) に到達するのは、さいころを  $4+3=7$  (回) 投げ、そのうち、1, 2, 3, 4 の目が 4 回出たときである.

よって、求める確率は、

$${}_7C_4 \left(\frac{4}{6}\right)^4 \left(\frac{2}{6}\right)^3 = {}_7C_3 \left(\frac{2}{3}\right)^4 \left(\frac{1}{3}\right)^3 = \frac{35 \times 2^4}{3^7} = \frac{560}{2187}$$

- (2) さいころを 200 回投げたとき、 $y$  の範囲は  $0 \leq y \leq 200$  より、確率  $P(0 \leq y \leq 200) = 1$  である.  
 よって、 $x$  座標が  $125 \leq x \leq 140$  である確率を求める.

200 回の試行で 1, 2, 3, 4 の目が出る回数を  $X$  とすると、

$$P(X=a) = {}_{200}C_a \left(\frac{2}{3}\right)^a \left(\frac{1}{3}\right)^{200-a} \quad (a=0, 1, 2, \dots, 200)$$

であるから、 $X$  の分布は二項分布  $B\left(200, \frac{2}{3}\right)$  であり、

$$Z = \frac{X - 200 \times \frac{2}{3}}{\sqrt{200 \times \frac{2}{3} \times \frac{1}{3}}} = \frac{3X - 400}{20}$$

とおくと、 $Z$  の分布は  $N(0, 1)$  とみなせる.

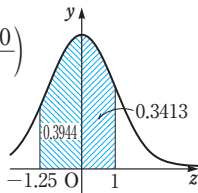
よって、

$$\begin{aligned} P(125 \leq X \leq 140) &= P\left(\frac{3 \cdot 125 - 400}{20} \leq Z \leq \frac{3 \cdot 140 - 400}{20}\right) \\ &= P(-1.25 \leq Z \leq 1) \\ &= P(0 \leq Z \leq 1.25) + P(0 \leq Z \leq 1) \\ &= 0.3944 + 0.3413 = 0.7357 \end{aligned}$$

したがって、求める確率は、**0.7357**

1, 2, 3, 4 の目が 4 回、  
5, 6 の目が 3 回  
出た場合である.

$$\begin{aligned} Z &= \frac{X - np}{\sqrt{npq}} \\ \text{において、} \\ n &= 200, \quad p = \frac{4}{6} = \frac{2}{3} \\ q &= 1 - p = \frac{1}{3} \end{aligned}$$



## 練習 B2.11

\*\*\*

3 個のさいころを 1 回投げたとき、3 個の目の和が 15 以上になる事象を  $E$  とする.

- (1) 事象  $E$  が起こる確率  $P(E)$  を求めよ.  
 (2) 3 個のさいころを 500 回投げるとき、事象  $E$  が少なくとも 60 回起こる確率を求めよ.



## Step Up

正規分布 ▶▶ 解答編 p.B2-23

第2章

\*\*\*

8

確率変数  $X$  が区間  $1 \leq x \leq 4$  の任意の値をとり、その確率密度関数は、←  
p.B2-20

$$f(x) = \begin{cases} -\frac{1}{2}(x-1)(x-3) & (1 \leq x \leq 2) \\ k(x-2) + \frac{1}{2} & (2 \leq x \leq 4) \end{cases} \quad (k \text{ は定数})$$

であるとき、次のものを求めよ。

- (1) 確率  $P(1 \leq X \leq 2)$                       (2)  $k$  の値  
(3) 平均  $m$  と分散  $V(X)$

\*\*\*\*

9

箱の中に1から  $n (\geq 2)$  までの数字が1つずつ書かれた  $n$  個の玉が入っている。この中から2個の玉を同時に取り出し、書かれている数の和だけポイントを受け取るゲームを行う。←  
p.B2-21

- (1) 受け取るポイント数  $X$  の平均  $m$  と標準偏差  $\sigma$  を求めよ。  
(2)  $n=123$  のとき、 $X \geq 155$  となる確率を求めよ。ただし、 $X$  は正規分布に従うものとし、 $\sqrt{186}=13.64$  とする。

\*\*

10

正規分布  $N(16, 10^2)$  に従う確率変数  $X$  について、次の等式が成り立つように、正の定数  $a$  の値をそれぞれ定めよ。←  
p.B2-22

- (1)  $P(X \geq a) = 0.8888$                       (2)  $P(12 \leq X \leq a) = 0.6342$   
(3)  $P(|X - 16| \geq a) = 0.4354$   
(4) 確率変数  $Y$  が正規分布  $N(0, 5^2)$  に従うとき、  
 $P(0 \leq Y \leq 6) = P(a \leq X \leq 18)$

\*\*\*

11

当たる確率の方が、はずれる確率よりも高くじ引きに、800名が参加したところ、当たる人数  $X$  の標準偏差は  $\frac{40}{3}$  であった。このとき、くじに当たる人が550人以上である確率を求めよ。

←  
p.B2-23

\*\*\*

12

袋の中に赤玉3個と白玉3個が入っている。この袋から同時に2個の玉を取り出す試行を900回行うとき、取り出した2個がともに赤玉である回数を  $X$  とする。

←  
p.B2-23確率  $P(X \geq k) \geq 0.75$  を満たす最大の整数  $k$  の値を求めよ。B1  
B2  
C1  
C2

## まとめ

## 3

## 統計的な推測

## 1. 母集団と標本

## 標本調査

統計調査には、調査の対象全体をもれなく調べる全数調査と、対象全体から一部を抜き出して調べ、それから全体を推測する標本調査の2通りの方法がある。

## 母集団と標本

標本調査では、調査の対象全体を母集団といい、調査のために母集団から抜き出された要素の全体を標本という。標本を抜き出すことを抽出といい、母集団、標本に含まれる要素の個数を、それぞれ母集団の大きさ、標本の大きさという。

## 無作為抽出と無作為標本

母集団の特徴を正しく推測するためには、標本にかたよりがなく、確率的にみて公平な抽出をしなければならない。そのために、乱数さい(特殊なさいころ)、乱数表、コンピュータで発生させた乱数を利用する。このような抽出方法を無作為抽出(または任意抽出)といい、無作為抽出によって選ばれた標本を無作為標本という。

## 復元抽出と非復元抽出

母集団から標本を抽出するとき、抽出のたびに要素をもとに戻し、あらためて次を抽出する方法を復元抽出という。一方、もとに戻さないで、続けて抽出する方法を非復元抽出という。

母集団から抽出された大きさ  $n$  の無作為標本  $(X_1, X_2, \dots, X_n)$  が、復元抽出によって得られたものであれば、 $X_1, X_2, \dots, X_n$  は独立である。また、非復元抽出によるものであれば、 $X_1, X_2, \dots, X_n$  は独立ではない。

ただし、母集団の要素の数がきわめて大きいときには、非復元抽出でも、 $X_1, X_2, \dots, X_n$  が独立であるとみなしてよいことが知られている。

## 母平均と標本平均

母集団分布についての平均を母平均、標準偏差を母標準偏差という。

母平均  $m$ 、母標準偏差  $\sigma$  の母集団から大きさ  $n$  の無作為標本  $(X_1, X_2, \dots, X_n)$  を抽出するとき、 $\frac{1}{n}(X_1 + X_2 + \dots + X_n)$  を標本平均といい、 $\bar{X}$  で表す。

また、標本平均  $\bar{X}$  の平均  $E(\bar{X})$  と標準偏差  $\sigma(\bar{X})$  は、

$$E(\bar{X}) = m, \quad \sigma(\bar{X}) = \frac{\sigma}{\sqrt{n}}$$

### 大数の法則

母平均  $m$  の母集団から、大きさ  $n$  の標本を無作為抽出するとき、 $n$  を大きくしていくと、標本平均  $\bar{X}$  は母平均  $m$  に近づく。

### 標本平均の標準化

母平均  $m$ 、母標準偏差  $\sigma$  の母集団から、大きさ  $n$  の標本を無作為抽出するとき、標本平均  $\bar{X}$  について、 $Z = \frac{\bar{X} - m}{\frac{\sigma}{\sqrt{n}}}$  は、 $n$  が大きいとき、ほぼ標準正規

分布  $N(0, 1)$  とみなすことができる。

【注】 母集団が正規分布に従うとき、 $n$  の大きさにかかわらず、標本平均  $\bar{X}$  は正規分布に従う。

第2章

## 2. 推定

与えられた標本から、母集団の平均や標準偏差など母集団の分布がもつ定数の値を推測することを推定という。

### 母平均の推定

標本の大きさ  $n$  が大きいとき、 $Z$  の値がある範囲にある確率は、正規分布表から求めることができる。

たとえば、 $P(|Z| \leq 1.96) = 0.4750 \times 2 = 0.95$  である。

$Z$  は、 $Z = \frac{\bar{X} - m}{\frac{\sigma}{\sqrt{n}}}$  であるから、 $|Z| \leq 1.96$  は、

$$|\bar{X} - m| \leq 1.96 \times \frac{\sigma}{\sqrt{n}}$$

$$\text{つまり、} \quad \bar{X} - 1.96 \times \frac{\sigma}{\sqrt{n}} \leq m \leq \bar{X} + 1.96 \times \frac{\sigma}{\sqrt{n}} \quad \cdots \cdots \textcircled{1}$$

となるから、①が成り立つ確率は 0.95 になる。

この①で示される範囲を、母平均  $m$  に対する信頼度 95% の信頼区間といい、

$\left[ \bar{X} - 1.96 \times \frac{\sigma}{\sqrt{n}}, \bar{X} + 1.96 \times \frac{\sigma}{\sqrt{n}} \right]$  で表す。また、信頼度 99% の信頼区間は、 $\left[ \bar{X} - 2.58 \times \frac{\sigma}{\sqrt{n}}, \bar{X} + 2.58 \times \frac{\sigma}{\sqrt{n}} \right]$  で表す。

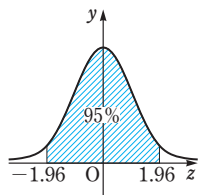
$\sigma$  の値がわからない場合、標本の大きさ  $n$  が大きいときには、 $\sigma$  を標本の標準偏差の値  $s$  とおき換えてもよい。さらに、次のことが知られている。

母平均  $m$  に対する信頼度 95% の信頼区間は、標本の大きさ  $n$  が大きいとき、標本平均の値を  $\bar{x}$ 、標本の標準偏差の値を  $s$  とすると、

$$\left[ \bar{x} - 1.96 \times \frac{s}{\sqrt{n}}, \bar{x} + 1.96 \times \frac{s}{\sqrt{n}} \right]$$

また、母平均  $m$  に対する信頼度 99% の信頼区間は、

$$\left[ \bar{x} - 2.58 \times \frac{s}{\sqrt{n}}, \bar{x} + 2.58 \times \frac{s}{\sqrt{n}} \right]$$



B1

B2

C1

C2

## 母比率の推定

母集団の中で、ある性質  $A$  をもつ要素の割合  $p$  を、その性質の母比率という。  
また、標本の要素のうち、性質  $A$  をもつ要素の割合を、その性質の標本比率という。

標本比率の値から、母集団の性質  $A$  をもつ要素の割合  $p$  を推測することを母比率の推定という。

母比率の推定の考え方は次のとおりである。

性質  $A$  をもてば 1、もたなければ 0 の値をとる確率変数を  $X$  とすると、その母平均は、 $E(X) = 0 \times (1-p) + 1 \times p = p$  であるから、母比率を推定するには、確率変数  $X$  の平均を推定すればよいことになる。

母集団から抽出した大きさ  $n$  の標本の性質  $A$  の標本比率を  $p_0$ 、標本平均を  $\bar{x}$  とすると、 $\bar{x} = 1 \times p_0 = p_0$ 、 $\bar{x}^2 = 1^2 \times p_0 = p_0$  より、標本の標準偏差  $s$  は、

$$s = \sqrt{\bar{x}^2 - (\bar{x})^2} = \sqrt{p_0 - p_0^2} = \sqrt{p_0(1-p_0)}$$

となる。

よって、 $n$  が大きいとき、母平均の推定と同様、母比率  $p$  に対する信頼度 95% の信頼区間は、

$$\left[ p_0 - 1.96 \times \sqrt{\frac{p_0(1-p_0)}{n}}, p_0 + 1.96 \times \sqrt{\frac{p_0(1-p_0)}{n}} \right]$$

同様に、信頼度 99% の信頼区間は、

$$\left[ p_0 - 2.58 \times \sqrt{\frac{p_0(1-p_0)}{n}}, p_0 + 2.58 \times \sqrt{\frac{p_0(1-p_0)}{n}} \right]$$

(例) 大量の製品 T から 100 個を無作為に抽出して調べたところ、8 個が不良品であった。

この製品 T の不良率  $p$  を信頼度 95% で推定すると、標本の大きさは、 $n=100$ 、

標本比率は、 $p_0 = \frac{8}{100} = 0.08$  より、標本の標準偏差は、

$$s = \sqrt{0.08(1-0.08)} \doteq 0.271$$

求める信頼区間は、

$$\left[ 0.08 - 1.96 \times \frac{0.271}{\sqrt{100}}, 0.08 + 1.96 \times \frac{0.271}{\sqrt{100}} \right]$$

すなわち、 $[0.027, 0.133]$  である。

### 3. 仮説検定

母集団に関する主張を数学的に記述したものを仮説という。

母集団に関する仮説を立て、それが正しいか否かを実験や観測に基づき判断する統計的手法を仮説検定という。

#### 帰無仮説と対立仮説

否定されることを想定した仮説を帰無仮説といい、 $H_0$  で表す。

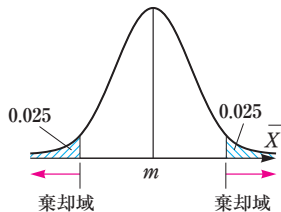
帰無仮説  $H_0$  を否定することで正当化されると考える仮説のことを対立仮説といい、 $H_1$  で表す。

仮説検定では、帰無仮説が否定できるかどうかを調べて、対立仮説を受け入れるかどうかを判断する。

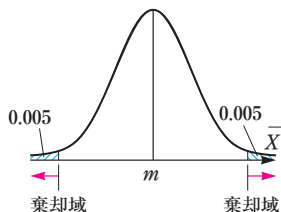
◀ 帰無仮説は、無に帰してはじめて意味が出てくる仮説である。

標本平均や標本比率について、帰無仮説を否定してもよいと考えられる値の範囲をあらかじめ定めておく必要がある。この範囲を棄却域という。棄却域は、基準となる数値  $\alpha$  をもとにして、標本平均や標本比率が棄却域に入る確率が  $\alpha$  となるように定める。この  $\alpha$  を有意水準または危険率といい、0.05 や 0.01 のように小さい値をとる。たとえば、 $\alpha=0.05$  のとき、その仮説検定を有意水準 5% の仮説検定という。

有意水準が 5% の場合



有意水準が 1% の場合



母集団から抽出した標本平均や標本比率の値が棄却域にあれば、帰無仮説は否定できると判断する。これを帰無仮説を棄却するという。このとき対立仮説を受け入れる。

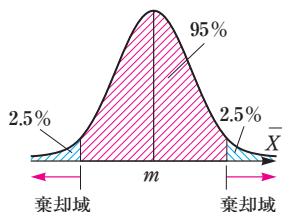
標本平均や標本比率の値が棄却域になければ、帰無仮説は否定できないと判断する。これを帰無仮説を棄却しないという。

【注】「帰無仮説を棄却しない」とは、帰無仮説を否定するだけの根拠が得られなかったという意味であり、帰無仮説が肯定されるということではない。

### 両側検定と片側検定

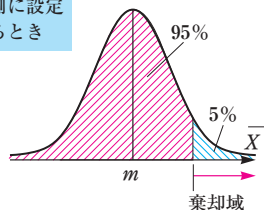
下の図のように棄却域を分布の両側に設定するような検定を両側検定という。

有意水準 5% の場合

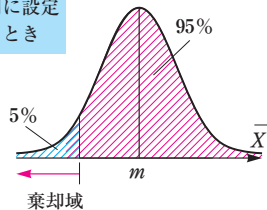


検定を行う際、棄却域を分布の片側に設定することもある。このような検定を片側検定という。

上側に設定  
するとき



下側に設定  
するとき



【注】 棄却域を上側に設定するときは  $m$  の右側の部分、下側に設定するときは  $m$  の左側の部分とする。

## Check!

\*\*

13

ある高校の2年生女子の身長は、平均が158.2 cm、標準偏差が6.5 cmの正規分布に従う。この中から9人を無作為抽出したとき、身長の平均が163 cm以上となる確率を求めよ。

\*

14

ある都市の商店街の客から無作為に100人を抽出して調べたところ、64人がその都市の住人であった。このとき、客全体の何%がその都市の住人であるかを、信頼度99%で推定せよ。

\*

15

ある菓子には内容量150gと記載されている。このことを仮説検定を用いて検討するとき、帰無仮説を述べよ。

\*\*

16

次の文章は、食品Xを製造しているある工場について述べたものである。次の(ア)～(ウ)のうち、仮説検定を用いて確かめることができないものをすべて選べ。

- (ア) ある工場が1日あたりに製造する食品Xの個数の平均値は300個である。
- (イ) ある工場で製造される食品Xは国内で1番価格が高い。
- (ウ) ある工場で1日に製造される食品Xの総数の99%は製品規格を満たす。

## 例題 B2.12 標本平均の平均・標準偏差

\*\*\*\*

ある都市での有権者の A 政党支持率は 40% である。この有権者の中から無作為に  $n$  人を抽出するとき、 $k$  番目の人が A 政党支持者なら 1 を、不支持者なら 0 の値を対応させる確率変数を  $X_k$  とし、標本平均を  $\bar{X}$  とする。

(1)  $\bar{X}$  の平均を求めよ。

(2)  $\bar{X}$  の標準偏差  $\sigma(\bar{X})$  が 0.04 以下となるための  $n$  の最小値を求めよ。

**考え方** 母集団から無作為に標本  $X_1, X_2, \dots, X_n$  を抽出すると、独立な確率変数  $X_1, X_2, \dots, X_n$  のそれぞれの平均  $E(X_k)$  と標準偏差  $\sigma(X_k)$  は、母集団と一致する。

**解答** (1) 母集団の確率分布は、A 政党支持なら 1、不支持なら 0 で、A 政党支持率は 40% より、右のようになる。

確率変数	1	0	計
確率	0.4	0.6	1

よって、母平均は、 $m=1 \times 0.4 + 0 \times 0.6 = 0.4$  より、 $E(X_k) = m = 0.4$

$$\begin{aligned}\bar{X} \text{ の平均は、} E(\bar{X}) &= E\left(\frac{1}{n}(X_1 + X_2 + \dots + X_n)\right) \\ &= \frac{1}{n}\{E(X_1) + E(X_2) + \dots + E(X_n)\}\end{aligned}$$

$$\begin{aligned}E(aX + bY) &= aE(X) + bE(Y) \\ E(X_1) &= E(X_2) = \dots \\ &= E(X_n) = m\end{aligned}$$

$$\text{よって、} E(\bar{X}) = \frac{1}{n}(m + m + \dots + m) = m = 0.4$$

(2) 母集団の標準偏差  $\sigma$  は、

$$\sigma = \sqrt{(1^2 \times 0.4 + 0^2 \times 0.6) - m^2} = \sqrt{0.4 - 0.4^2} = \sqrt{0.24}$$

であり、標本平均  $\bar{X}$  の標準偏差は、

$$\begin{aligned}\sigma(\bar{X}) &= \sqrt{V(\bar{X})} = \sqrt{V\left(\frac{1}{n}(X_1 + X_2 + \dots + X_n)\right)} \\ &= \sqrt{\frac{1}{n^2}\{V(X_1) + V(X_2) + \dots + V(X_n)\}} \\ &= \sqrt{\frac{1}{n^2}(\sigma^2 + \sigma^2 + \dots + \sigma^2)} = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}\end{aligned}$$

$X_1, X_2, \dots, X_n$  は独立とみなしてよい。

$$\begin{aligned}X, Y \text{ が独立のとき、} \\ V(aX + bY) &= a^2V(X) + b^2V(Y)\end{aligned}$$

$$\text{したがって、} \sigma(\bar{X}) = \frac{\sqrt{0.24}}{\sqrt{n}} \text{ より、} \frac{\sqrt{0.24}}{\sqrt{n}} \leq 0.04 \text{ であるから、}$$

$$\sqrt{n} \geq \frac{\sqrt{0.24}}{0.04} \text{ より、} n \geq \frac{0.24}{0.0016} = 150 \quad \text{よって、} n \text{ の最小値は } 150$$

## Focus

母平均  $m$ 、母標準偏差  $\sigma$  の母集団から、大きさ  $n$  の標本を無作為に抽出するとき、標本平均  $\bar{X}$  の

$$\text{平均(期待値) } E(\bar{X}) = m \quad \text{標準偏差 } \sigma(\bar{X}) = \frac{\sigma}{\sqrt{n}}$$

練習  
B2.12  
\*\*\*

右のような確率分布の母集団から、大きさ 100 の標本を無作為に抽出する。その標本平均  $\bar{X}$  の平均  $E(\bar{X})$  と標準偏差  $\sigma(\bar{X})$  を求めよ。

→ p.B2-42 [13]

$X$	1	2	3	4	計
$P$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$	1



## 例題 B2.13 標本平均の標準化

\*\*\*\*

大勢の生徒が受験した、ある試験の得点  $X$  は正規分布  $N(240, 28^2)$  に従っている。この母集団から 25 人を無作為に抽出するとき、25 人の得点の平均  $\bar{X}$  が  $230 \leq \bar{X} \leq 250$  となる確率を求めよ。

**考え方** 母平均  $m$ 、母標準偏差  $\sigma$  の母集団から大きさ  $n$  の標本を無作為抽出すると、標本平均  $\bar{X}$  の平均(期待値)は、 $E(\bar{X}) = m$ 、標本平均  $\bar{X}$  の標準偏差は、 $\sigma(\bar{X}) = \frac{\sigma}{\sqrt{n}}$  である。  
 $X$  は正規分布に従うから、 $Z = \frac{\bar{X} - m}{\frac{\sigma}{\sqrt{n}}}$  とおくと、 $Z$  の確率分布は標準正規分布  $N(0, 1)$

に従う。

**解答** 母集団の得点  $X$  は、平均  $m=240$ 、標準偏差  $\sigma=28$  の正規分布に従うから、大きさ 25 の標本の標本平均  $\bar{X}$  の平均は、

$$E(\bar{X}) = 240$$

標準偏差は、

$$\sigma(\bar{X}) = \frac{28}{\sqrt{25}} = \frac{28}{5}$$

よって、

$$Z = \frac{\bar{X} - 240}{\frac{28}{5}} = \frac{5\bar{X} - 1200}{28}$$

とおくと、 $Z$  は標準正規分布  $N(0, 1)$  に従うから、 $230 \leq \bar{X} \leq 250$  より、

$$\frac{5 \times 230 - 1200}{28} \leq Z \leq \frac{5 \times 250 - 1200}{28}$$

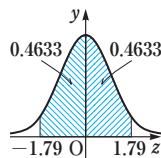
つまり、 $-1.79 \leq Z \leq 1.79$

よって、求める確率は、

$$\begin{aligned} & P(-1.79 \leq Z \leq 1.79) \\ &= P(-1.79 \leq Z \leq 0) + P(0 \leq Z \leq 1.79) \\ &= 2 \times P(0 \leq Z \leq 1.79) \\ &= 2 \times 0.4633 \\ &= 0.9266 \end{aligned}$$

$$E(\bar{X}) = m = 240$$

$$\begin{aligned} \sigma(\bar{X}) &= \frac{\sigma}{\sqrt{n}} \\ &= \frac{28}{\sqrt{25}} \end{aligned}$$



**注** 同じ母集団からの標本で、標本平均  $\bar{X}$  の範囲が同じであっても、その確率は、標本の大きさ  $n$  に左右される。

## 練習 B2.13

例題 B2.13 と同じ母集団から、 $n$  人を無作為に抽出するとき、 $n$  人の得点の平均  $\bar{X}$  が  $230 \leq \bar{X} \leq 250$  となる確率が 0.95 以上となる  $n$  の最小値を求めよ。

\*\*

→ p. B2-42 14

## 例題 B2.14 母平均の推定

\*\*\*\*

ある高校2年生の男子の中から無作為に抽出した100人の身長は下のようであった。この高校2年生の男子の平均身長を信頼度95%で推定せよ。ただし、 $\sqrt{555}=23.6$ として計算せよ。

身長	以上 未満	150 155	155 160	160 165	165 170	170 175	175 180	180 185	計
人数		1	4	17	35	26	14	3	100

**考え方** 母標準偏差  $\sigma$  がわからない場合、標本の大きさ  $n$  が大きいときは、標本の標準偏差  $s$  を用いても差し支えない。そこで、与えられたデータから、標本の標準偏差  $s$  を求める。

**解答** 右の表は、階級値  $x$  ごとに、度数  $f$ 、階級値 167.5 を仮平均としたと

きの  $y = \frac{x-167.5}{5}$  の値、

また、 $yf$ 、 $y^2f$  の値とその縦の合計をまとめたものである。

$x=5y+167.5$  であるから、標本平均は、

$$\bar{x} = E(x) = E(5y + 167.5) = 5E(y) + 167.5 = 5 \times \frac{35}{100} + 167.5 = 169.25$$

$$\text{標本の標準偏差は、} s = 5 \sqrt{\frac{151}{100} - \left(\frac{35}{100}\right)^2} = \frac{\sqrt{555}}{4}$$

$$\text{標本の大きさ } 100, \text{ 標本平均 } 169.25, \text{ 標本の標準偏差 } \frac{\sqrt{555}}{4} \quad \left| \begin{array}{l} a, b \text{ が定数で,} \\ x = ay + b \text{ のとき,} \\ \sigma(x) = |a|\sigma(y) \end{array} \right.$$

より、この高校2年生男子の平均身長に対する信頼度95%の信頼区間は、

$$\left[ 169.25 - 1.96 \times \frac{\sqrt{555}}{4} \times \frac{1}{\sqrt{100}}, 169.25 + 1.96 \times \frac{\sqrt{555}}{4} \times \frac{1}{\sqrt{100}} \right]$$

つまり、 $[168.1, 170.4]$

階級値のままでは計算が大変なので、 $y = \frac{x-167.5}{5}$  といっておく。

## Focus

標本の大きさ  $n$  が大きいとき、標本平均の値を  $\bar{x}$ 、標本の標準偏差の値を  $s$  とすると、母平均に対する信頼度95%の信頼区間は、

$$\left[ \bar{x} - 1.96 \times \frac{s}{\sqrt{n}}, \bar{x} + 1.96 \times \frac{s}{\sqrt{n}} \right]$$

## 練習 B2.14

\*\*

ある高校で50人の生徒を無作為に抽出し、5月の読書冊数を調べたところ、下の表のようになった。この高校における、1人当たりの5月の読書冊数の平均を、信頼度99%で推定せよ。ただし、 $\sqrt{330}=18.2$ として計算せよ。

読書冊数	0	1	2	3	4	5	計
人数	8	18	12	7	3	2	50

→ p.B2-42 15

## 例題 B2.15 母比率の推定

\*\*\*\*

ある新製品から 225 個を無作為抽出して調べたところ、6% が不良品であった。この新製品の不良品率(母比率)を信頼度 95% で推定せよ。

ただし、 $\sqrt{0.0564}=0.24$  として計算せよ。

第2章

考え方

標本比率  $p_0$  から、標本の標準偏差  $s=\sqrt{p_0(1-p_0)}$  を求める。

標本の大きさ  $n$  が大きいとき、母比率に対する信頼度 95% の信頼区間は、

$$\left[ p_0 - 1.96 \times \frac{s}{\sqrt{n}}, p_0 + 1.96 \times \frac{s}{\sqrt{n}} \right]$$

である。

解答

標本比率は  $p_0 = \frac{6}{100} = 0.06$  であるから、標本の標準偏差は、

$$\begin{aligned} s &= \sqrt{p_0(1-p_0)} \\ &= \sqrt{0.06 \times 0.94} \\ &= \sqrt{0.0564} = 0.24 \end{aligned}$$

よって、標本の大きさが 225 より、母比率に対する信頼度 95% の信頼区間は、

$$\left[ 0.06 - 1.96 \times \frac{0.24}{\sqrt{225}}, 0.06 + 1.96 \times \frac{0.24}{\sqrt{225}} \right]$$

すなわち、 $[0.029, 0.091]$

$$\begin{aligned} &1.96 \times \frac{0.24}{\sqrt{225}} \\ &= 1.96 \times \frac{0.24}{15} \\ &= 0.03136 \end{aligned}$$

## Focus

標本の大きさ  $n$  が大きいとき、標本比率を  $p_0$  とすると、母比率に対する信頼度 95% の信頼区間は、

$$\left[ p_0 - 1.96 \times \sqrt{\frac{p_0(1-p_0)}{n}}, p_0 + 1.96 \times \sqrt{\frac{p_0(1-p_0)}{n}} \right]$$

【注】 99% の信頼区間の場合、1.96 のかわりに 2.58 を掛ける。

## 練習 B2.15

\*\*

ある高校の 3 年生の生徒から 150 人を無作為抽出して調べたところ、理系の生徒は 96 人であった。この高校の 3 年生全体における理系の生徒の割合を信頼度 99% で推定せよ。

ただし、 $\sqrt{6}=2.45$  として計算せよ。

→ p.B2-42 (16)

B1  
B2  
C1  
C2

## Think

## 例題 B2.16 仮説検定(両側検定)

\*\*\*\*

ある果物 A の原産地表示が国内産となっているが、外国産ではないかという疑惑が生じた。そこで、100 個の標本を無作為抽出し、その大きさを調べたところ、その平均は 5.8 cm であった。果物 A が国内産である場合は、その大きさの母平均は 6.0 cm、標準偏差は 1.0 cm であり、外国産である場合は、その大きさの母平均は 6.0 cm ではないことがわかっている。この果物 A が国内産であるといえるかを仮説検定するとき、次の問いに答えよ。

- (1) 帰無仮説  $H_0$  を「果物 A の大きさの母平均は 6.0 cm である」としたとき、対立仮説  $H_1$  を述べよ。
- (2) 果物 A が国内産であるといえるかを有意水準 5% で仮説検定せよ。

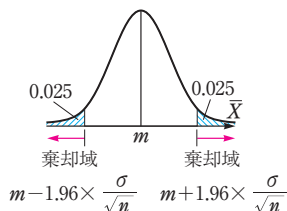
**考え方** 仮説検定を行う際には、次の手順で進めていけばよい。

**【仮説検定の手順】**

- (I) 仮説検定を行って示したい仮説  $H_1$  (対立仮説) と、否定されることで対立仮説の正しさを示す帰無仮説  $H_0$  を立てる。
- (II) 有意水準  $\alpha$  を定める。
- (III) 適切な検定統計量  $T$  を選び、実験や観測のデータから、その現実値  $t$  を求める。
- (IV) 帰無仮説の下で、検定統計量  $T$  が現実値  $t$  より極端な値が生じる確率  $P$  を求める。
  - (i)  $P < \alpha$  の場合、帰無仮説  $H_0$  を棄却する。
  - (ii)  $P \geq \alpha$  の場合、帰無仮説  $H_0$  を棄却しない。

果物 A が外国産であるという疑惑があるので、対立仮説としては、「果物 A の大きさの母平均が 6 cm ではない」となる。果物の大きさの母平均を  $m$  cm とすると、対立仮説  $H_1$  は、 $H_1: m \neq 6$  となる。

この問題では、外国産の果物であった場合、標本平均が 6 cm よりも大きいかわかりませんが、外国産が 6 cm よりも小さい」というようなことがわかっている場合は、片側検定で行う（注）参照）。



次に、(II)の有意水準を定める。この場合は、5% である。

母標準偏差  $\sigma$  の母集団において、帰無仮説として母平均が  $m$  であると仮定して両側検定を行うとき、その母集団から無作為抽出した大きさ  $n$  の標本の標本平均  $\bar{X}$  についての有意水準 5% の棄却域は、

$$|\bar{X} - m| > 1.96 \times \frac{\sigma}{\sqrt{n}}$$

で与えられる（p.B2-41 参照）。

次に、(III)のデータであるが、この場合は、標本 100 個のデータがあるのでそれを用いる。そのデータから、検定統計量  $T$  が現実値以上に極端な値をとる確率  $P$  を求め、有意水準 5% と比較する。その際、上で考えた棄却域を利用することができる。

## 解答

- (1) 帰無仮説は、「果物 A の大きさの母平均は 6.0 cm である」なので、対立仮説は、  
 「果物 A の大きさの母平均は 6.0 cm ではない」である。  
 (2) 母平均を  $m$ 、母標準偏差を  $\sigma$ 、標本の大きさを  $n$ 、標本平均を  $\bar{X}$  とすると、 $\bar{X}$  の棄却域は、

$$|\bar{X} - m| > 1.96 \times \frac{\sigma}{\sqrt{n}} \quad \cdots \cdots \textcircled{1}$$

$\sigma = 1.0$ 、 $n = 100$  であり、帰無仮説より、 $m = 6.0$  である。

これらを①に代入すると、棄却域は、

$$|\bar{X} - 6.0| > 1.96 \times \frac{1.0}{\sqrt{100}} = 0.196 \text{ を満たす範囲である。}$$

ここで、 $\bar{X} = 5.8$  とすると、 $|5.8 - 6.0| = 0.2 > 0.196$  となり、帰無仮説は棄却される。

よって、果物 A の大きさの母平均は 6.0 cm でないといえるので、果物 A は国内産でないといえる。

帰無仮説を棄却する

↓

対立仮説は正しいと判断

**注** 例題 B2.16 では、果物 A が外国産の場合に、母平均は 6.0 cm ではないことはわかっているが、6.0 cm より大きいかわかりか小さいかわからないため、母平均は 6.0 cm ではないことを対立仮説として、両側検定を用いた。もし、「外国産の場合は、6 cm よりも小さいことがわかっている」という場合は、次のように片側検定を用いるとよい。

帰無仮説：「果物 A の大きさの母平均は 6 cm である」

対立仮説：「果物 A の大きさの母平均は 6 cm より小さくなる」

片側検定の場合は、有意水準 5% の棄却域は、分布の右側もしくは左側だけに設定する。この場合は、左側だけに設定する。

このとき、棄却域は、

$$\bar{X} - m < -1.64 \times \frac{\sigma}{\sqrt{n}}$$

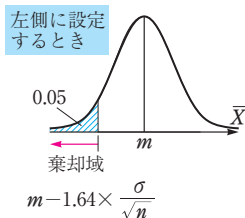
で与えられる ( $p$ .B2-41 参照)。

ここで、 $\sigma = 1.0$ 、 $n = 100$ 、 $m = 6.0$ 、 $\bar{X} = 5.8$  とすると、

$$\bar{X} - m = 5.8 - 6.0 = -0.2, \quad -1.64 \times \frac{\sigma}{\sqrt{n}} = -1.64 \times \frac{1.0}{\sqrt{100}} = -0.164 \text{ であり、} -0.2 < -0.164$$

であるから、帰無仮説は棄却される。

よって、果物は国内産でないといえる。



練習  
B2.16

\*\*\*

ある店で売られているチョコレートは1個の重さが10gであるという。このチョコレートを196個購入し、重さを調べたところ、標本平均は10.05g、標準偏差は0.04gであった。この店のチョコレート1個の重さの平均は10gといえるか、有意水準5%で仮説検定せよ。

B1  
B2  
C1  
C2

## Story ストーリー

## 「仮説検定の考え方」

次の問題を考えてみよう。

【問題 1】A 君は走り高跳びの選手で、今までは 180 cm の高さを跳べる確率は  $\frac{1}{2}$  であった。しかし、トレーニング法を変えてみたところ、実力が上がったような気がした。

そこで、顧問の先生にそのことを伝えたところ、先生が「今から 8 回跳んで 7 回以上成功したら実力向上を認める」と言った。実際に A 君は 8 回中 7 回成功し、実力向上を認めてもらった。

これを仮説検定の枠組みで考えると、

帰無仮説  $H_0$  : 成功確率  $p = \frac{1}{2}$  (実力は向上していない)

対立仮説  $H_1$  : 成功確率  $p > \frac{1}{2}$  (実力は向上している)

帰無仮説  $H_0$  を仮定して、8 回跳んで 7 回以上成功する確率は、反復試行の確率より、

$$\left(\frac{1}{2}\right)^8 + {}_8C_1 \left(\frac{1}{2}\right)^7 \cdot \frac{1}{2} = \frac{9}{256} \div 0.035 = 3.5\%$$

と明らかに低い、よって、帰無仮説  $H_0$  を仮定したことが誤りで、帰無仮説  $H_0$  の否定である対立仮説  $H_1$  を採択する。つまり、実力向上を認めるということである。

このように、帰無仮説は否定（小さい確率のことが起こってしまう）されることによって、対立仮説を採択する。つまり、無に帰してはじめて意味が出てくる仮説（仮定）なので、「帰無仮説」と呼ばれる。

ただ、上の論議での問題点はある。

どれくらい小さい確率で帰無仮説を棄却（否定）するのかで、通常であれば有意水準と呼ばれる確率 5% を設定し、議論を行う（有意水準が 10%、1% の場合もよくある設定である※ 1）。

さらに、次のような問題を考えてみよう。

【問題 2】高跳びの跳ぶ回数が増えれば増えるほど成功確率（の推定）は安定してくるので、顧問の先生が「80 回のうち 50 回以上成功したら実力向上を認める。」と言った。  
これについて調べてみよう。

帰無仮説  $H_0$  : 成功確率  $p = \frac{1}{2}$  (実力は向上していない)

対立仮説  $H_1$  : 成功確率  $p > \frac{1}{2}$  (実力は向上している)

のどちらが正しいかを統計的に調べる。

帰無仮説  $H_0$  の仮定の下では、成功確率は  $\frac{1}{2}$  であるから、80 回のうち成功回数を  $X$  とすると、 $X$  は二項分布  $B\left(80, \frac{1}{2}\right)$  に従う。

すると、 $E(X) = 80 \times \frac{1}{2} = 40$ 、 $V(X) = 80 \times \frac{1}{2} \times \frac{1}{2} = 20$  であり、  
 $Z = \frac{X - E(X)}{\sqrt{V(X)}} = \frac{X - 40}{\sqrt{20}}$  とすると、 $Z$  は近似的に標準正規分布に従う、したがって、

$$\begin{aligned} P(X \geq 50) &= P\left(\frac{X - 40}{\sqrt{20}} \geq \frac{50 - 40}{\sqrt{20}}\right) \\ &= P(Z \geq \sqrt{5}) \\ &\doteq P(Z \geq 2.24) \\ &= 0.5 - 0.4875 = 0.0125 \end{aligned}$$

と明らかに小さいので、帰無仮説を棄却して対立仮説を採択する。

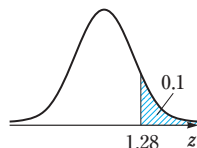
ここで、有意水準を 5% と定めると、 $P(Z \geq 1.64) = 0.5 - 0.4495 \doteq 0.05$  であるから、

$$\begin{aligned} P\left(\frac{X - 40}{\sqrt{20}} \geq 1.64\right) &= P(X \geq 40 + 1.64 \times 2\sqrt{5}) \\ &\doteq P(X \geq 47.33) \end{aligned}$$

つまり、80 回のうち 48 回以上跳ぶことができたなら有意水準 5% で実力向上を認定（採択）できる。

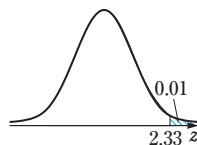
有意水準が 10% なら,  $P(Z \geq 1.28) = 0.5 - 0.3997 \div 0.1$

$P\left(\frac{X-40}{\sqrt{20}} \geq 1.28\right) \div P(X \geq 45.72)$  より, 80 回のうち 46 回以上跳ぶことができたなら実力向上を認定する.



有意水準が 1% なら,  $P(Z \geq 2.33) = 0.5 - 0.4901 \div 0.01$

$P\left(\frac{X-40}{\sqrt{20}} \geq 2.33\right) \div P(X \geq 50.42)$  より, 80 回のうち 51 回以上跳ぶことができたなら実力向上を認定する.



このように, 有意水準を何%に設定するかによって, 実力向上と認定できる回数は異なるが, 信頼度がどれくらいをもって検定したいかを考えて有意水準を設定すればよい.

また, 「8 回跳んで 7 回以上で実力向上を認定する」ことから, 「80 回跳んで 70 回以上で実力向上を認定する」と単純に 10 倍してはいけないことにも注意が必要である.

#### ※ 1

有意水準 1% は, 信頼度 99% であるから, 最も高い水準となる.

同様に, 有意水準 5% は信頼度 95%, 有意水準 10% は信頼度 90% となる.

有意水準 1%, 5%, 10% と設定することが多いが, これは検定を行う際に, どれくらい信頼度をもつか考えて判断すればよい.

たとえば, 有意水準を 50% としてしまうと, 信頼度も 50% となり, 検定を行う意味がないことがわかる.



## Column コラム

## 「有意水準と棄却域について」

例題 B2.16 で両側検定を行ったが、その際に用いた不等式について考えてみよう。  
まず、両側検定について確認しておこう。

母平均  $m$ 、母標準偏差  $\sigma$  の母集団から無作為抽出した大きさ  $n$  の標本の標本平均  $\bar{X}$  について、

$$|\bar{X} - m| \leq 1.96 \times \frac{\sigma}{\sqrt{n}}$$

が成り立つ確率は 0.95 である。このことから、帰無仮説として母平均が  $m$  であると仮定したとき、標本平均  $\bar{X}$  についての有意水準 5% の棄却域は

$$|\bar{X} - m| > 1.96 \times \frac{\sigma}{\sqrt{n}}$$

を満たす範囲となる。このように、棄却域を分布の両側に設定するような検定を両側検定といった。このときの、「1.96」について考えてみよう。

母平均  $m$ 、母標準偏差  $\sigma$  の母集団から無作為抽出した大きさ  $n$  の標本の標本平均  $\bar{X}$  は、 $n$  が大きいとき、正規分布  $N\left(m, \frac{\sigma^2}{n}\right)$  に従うとしてよく、

$Z = \frac{\bar{X} - m}{\frac{\sigma}{\sqrt{n}}}$  は、標準正規分布  $N(0, 1)$  に従うとしてよい。

正規分布表を調べると、 $P(|Z| \leq 1.96) = 0.95$  であるから、

$$P\left(|\bar{X} - m| \leq 1.96 \times \frac{\sigma}{\sqrt{n}}\right) = 0.95$$

となり、ここから信頼度 95% の信頼区間が得られる。信頼度 95% となるので、有意水準 5% の棄却域が設定できるのである。

同様に、片側検定を行ったときの棄却域について考えてみよう。

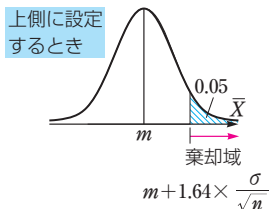
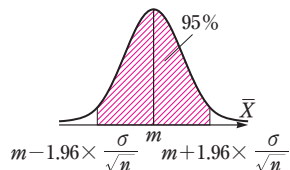
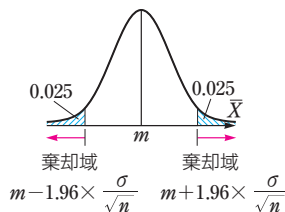
有意水準 5% のとき、 $|\bar{X} - m| > 1.64 \times \frac{\sigma}{\sqrt{n}}$  ……①

を満たす範囲のうち、棄却域を上側に設定するときは、 $m$  の右側、下側に設定するときは、 $m$  の左側の部分とする。

たとえば、上側に設定するとき、 $P(0 \leq Z \leq x) = 0.45$  を満たす  $x$  は正規分布表より、 $x = 1.64$  のとき最も近い値を示す。このようにして、①が得られる。

下側に設定する場合は、 $P(-1.64 \leq Z \leq 0) = 0.45$  となる。

有意水準は 5% 以外にも 1% や 10% を設定する場合があるが、そのときの棄却域も、同様に考えればよい。ここでは、棄却域を示す式を記しておくので、ぜひ各自で確認しておいてほしい。



	両側検定のときの棄却域	片側検定（上側）のときの棄却域
有意水準 1%	$ \bar{X} - m  > 2.58 \times \frac{\sigma}{\sqrt{n}}$	$\bar{X} - m > 2.33 \times \frac{\sigma}{\sqrt{n}}$
有意水準 10%	$ \bar{X} - m  > 1.64 \times \frac{\sigma}{\sqrt{n}}$	$\bar{X} - m > 1.28 \times \frac{\sigma}{\sqrt{n}}$

## Step Up

統計的な推測 ▶▶ 解答編 p.B2-36

\*\*\*

13

←  
p.B2-32さいころを  $n$  回投げて、出た目の表す確率変数を順に  $X_1, X_2, \dots, X_n$ とする.  $\bar{X} = \frac{1}{n} \sum_{k=1}^n X_k$  とするとき,

- (1)  $\bar{X}$  の期待値  $E(\bar{X})$  と分散  $V(\bar{X})$  を求めよ.
- (2)  $n=3$  のとき,  $|\bar{X} - E(\bar{X})| \geq 2\sqrt{V(\bar{X})}$  となる確率を求めよ. (九州大)

\*\*\*

14

←  
p.B2-33平均が 12.5, 標準偏差が 4.2 で, 正規分布に従う数の集合  $A$  がある. この集合のすべての数を 4 倍して 5 を加えて新たな集合  $B$  を作ることにする.

- (1) 集合  $B$  に属する数の平均と標準偏差を求めよ.
- (2) 集合  $B$  を母集団として, 大きさ 64 の標本を抽出する. その標本の平均が 54 より小さくなる確率を求めよ.

\*\*\*

15

←  
p.B2-34ある農業生産物の長さの分布は, 母平均が  $m$  cm, 母標準偏差が 1.8 cm の正規分布に従っている. 大きさ  $n$  の標本を無作為に抽出し, 母平均  $m$  に対する信頼度 95% の信頼区間を調べたところ,  $[23.696, 24.704]$  であった.

- (1) 標本平均  $\bar{x}$  と  $n$  の値を求めよ.
- (2) 母平均  $m$  に対する信頼度 99% の信頼区間を推定せよ.

\*\*\*

16

←  
p.B2-35ある地区の 100 組の夫婦を対象に, A 製品, B 製品のどちらを選ぶかを調査した. その結果, 夫の A 製品, B 製品の選択者数はそれぞれ 60 名, 40 名で, 妻の A 製品, B 製品の選択者数はそれぞれ 75 名, 25 名であった. この地区全体について, 夫婦の選択が一致する確率を信頼度 95% で推定せよ. なお, 夫と妻が A, B いずれの製品を選択するかの確率は互いに独立であるとする. また,  $\sqrt{11} = 3.32$  として計算せよ.

## 章末問題

▶▶ 解答編 p.B2-40

\*\*\*

1

←  
p.B2-34

−1, 0, 1, 2, 3 の数が書かれている札が, それぞれ 1 枚, 2 枚, 4 枚, 3 枚, 2 枚ずつ箱に入っている. これを母集団とし, 箱の中から無作為に 1 枚を抽出しては箱に戻す試行を  $n$  回繰り返す. 札に書かれた数を確率変数  $X$  とするとき,

- (1) 母平均と母標準偏差を求めよ.
- (2)  $k$  回目 ( $k=1, 2, \dots, n$ ) に取り出した札の数を  $X_k$  とする.  
母平均  $m$  に対する信頼度 99% の信頼区間の幅を 0.8 以下で推定するには, 試行回数  $n$  をどの程度の大きさとすればよいか.

第2章

## 思考力問題

\*\*\*\*

2

$n$  を自然数として 1 枚のコイン投げを  $2n$  回行う. この  $2n$  回のコイン投げで, 表が出る合計回数を  $X$  とする. ただし, コインの表と裏の出る確率は等しいとする. 次の各問いに答えよ.

- (1)  $X$  の期待値と標準偏差をそれぞれ求めよ.
- (2)  $\frac{P(X=k+1)}{P(X=k)}$  を求めよ. ただし,  $k=0, 1, 2, \dots, 2n-1$  とする.
- (3)  $P(X=k)$  を最大にする  $k$  の値を求めよ.
- (4)  $n=200$  とする. 試行回数が大きいとき,  $X$  の確率分布は正規分布で近似できることが知られており, 試行回数 400 はこのような近似が成り立つのに十分大きいとみなせる. このことを利用して,  $X$  の値が

$$190 \leq X \leq 210$$

となる確率の近似値を求めよ. ただし, 標準正規分布に従う確率変数  $Z$  に対する  $P(Z>1)$  の近似値としては 0.159 を用いよ.

(鹿兒島大)

B1

B2

C1

C2

## Coffee Break

コーヒーブレイク

### 「著者からのメッセージ①」

日本では、東日本大震災や熊本地震など様々な災害が起っています。また、復興に向けての様々な取り組みがされているもののまだまだ多くの方々が大変な思いをされています。そんな中でも、子供たち（児童、生徒）にとって、まだまだ安心して学べる環境が整っていないところが多々あります。子ども達（あなた方）は未来の宝です。彼らが一日も早く明るい笑顔で、再び良い仲間と遊び、学ぶ環境が取り戻せることを心より願っています。

今回、なぜこのような話を取り上げたかという、我々は普段の生活の中では当たり前すぎて気づかないことが多いという話をしたかったからです。毎日食べている食物、電気、水道、住居、そして家族、友達など、自分の周りにあるのが当然と思っていませんか。これは、「教育」という点についても同じです。我々は当たり前のように小中学校という義務教育課程に進み、またそのほとんどが高校に進みます。学校では教科書が配られ、先生達が生徒のために丁寧に授業をしてくれる。これがどれほど当たり前でないかは、広い世界を見渡せば明らかです。世界の多くでは、学校自体が存在しない、学校に通いたくても通えない、鉛筆1本が買えない、1冊の教科書をクラスの全員で使う……。彼らの多くは、日々生きていくのが精一杯で、勉強どころではないのです。つまり勉強したくても勉強できない環境に置かれています。しかし、彼らの「学びたい」という意欲は旺盛です。一旦、学ぶ機会が与えられれば、目を輝かせて何でもかんでも貪欲に吸収しようとする姿があります。それに比べて我々はどうでしょう。〇〇だからやる気が起きない、〇〇だから数学ができるようにならない、塾に行かないと〇〇大学には受からない。

自分がどれほど恵まれた環境にいるかも知らずに「もっと、もっと」と求め続けている。この欲求は、「彼らの学びの欲求」とは大きく違うような気がします。

ここで皆さんにもう一度考えて欲しいと思います。

「学ぶとはどういうことか」、「自分は何のために学びたいのか」、そして「自分が恵まれた環境の中で学べる幸せ」ということを。

最後に皆さんに送りたい文章があります。これは僕が高校の教師をしていた頃からよく生徒達に贈っていた文で、著者の丹羽健夫さんに快諾いただき、これまでも全国の皆さんに届けてきました。今回はその文章を Coffee Break を通して皆さんにお伝えしたいと思います。心より感謝致します。この文章から何かを感じ取っていただければと思います。

浪人して予備校にやってくるのは、何も高校出たての一浪生、二浪生ばかりとは限らない。小野道風よろしく初志を貫徹するため、A大学医学部をめざして五浪中というような人もいる。そしてもっと年齢のいった二十代後半とか三十代の人も少数ではあるが。一旦大学を卒業して社会に出たけれど、実務の中で「俺が(私が)本当にやりたかったのは××だったのだ」と人生的発見をして改めて別の進路をやり直すために予備校へ通ってくる人たちが、数年前のこと。四月の開講日に、予備校の教室に一人の年配の男性が現れた。チューター(クラス担任)が、「保護者の方ですか」と聞くとその男性はニヤッと笑って、「いえ、れっきとした塾生です。かなり年がいつているので、ご迷惑をおかけするかもしれませんがよろしくお願いします」と息子ぐらいの年配のチューターに丁寧に頭を下げた。翌日からの授業にそのお年寄りは、授業開始三〇分前ぐらいに、紫色の風呂敷にテキストやノートを包んで教室に現れ、最前列に座るのがつねとなった。講師たちの間でも「オイ、七〇一教室の一番前にへんなオッサンがおるな」と噂になった。ひと月ほどしてチューター面談のときにこのお年寄りの状況が少し明かになった。年齢は六十一歳。前の年に会社を定年になったのであった。「私が旧制中学を卒業したのは、太平洋戦争で日本が敗れて一〇年もたってないときでした。私らの一級下から新制中学、新制高校になったのです。当時は日本中が貧乏で、ましてや私は父親が戦死していました。母親が勤めに出て何とか中学までは行けましたけど、どうして旧制の高校、大学には行けませんでした。割合勉強もできる方だったんですが、母親が苦勞しているのをみて少しでも早く勤めに出て安心させたかったんです。もちろん私はできれば進学したかったんですが——。就職するときに決心したのです。勤めに出て少しでもお金がたったら大学はそれからでも遅くないと。しかし、一旦就職すると仕事に追われて時間的にそれどころではなく、そのうちに結婚して子どもができて、そうなる経済的に大変で大学どころではなくなりました。昨年やっと仕事で定年になりまして、子どもたちもそれぞれに道を歩みはじめて少しばかりのお金も残りましたので、こうしてやっと本式に大学へ向けての勉強にとりかかることができたのです。ええ、女房もそりゃあ喜んでくれています……」年のせいと顔が錆びついたり、などとこぼしながらもお年寄り是一日も欠かさずきちんと予習をして授業に出席した。講師たちも「あのお年寄りを絶対ものにしようや」と力を貸した。塾生たちの間でも「おとうさん」といって慕われ、体育大会でも若いものに混じって二人三脚やムカデ競争に出場した。成績の方は夏を過ぎる頃から急伸をはじめた。十一月の受験校決定のときチューターは、第一志望の、地元の私立大で一番の難関のB大学文学部史学科はまず間違いなしと太鼓判を押した。B大学の合格発表のあった二

月半ば、チューターは真っ先にお年寄りの名前を探した。しかしお年寄りの名前はなかった。チューターは焦って家に電話した。「いやあ、実は受けませんでしてな。どこの大学も受けませんでしてな。考えが変わったんです。もし私が大学に受かったら、一緒に勉強しておった若い人たちの席を、一つ奪うことになる気がつきましてな。こちらは人生の残りで大学に行こうとしとるんだが、若い人たちにとっては、これからの人生のために大学に行くんですからな。いやあ、一年間楽しく勉強させていただいて十分満足ですわ。合格祝賀会には行って皆の顔をみます。ええ、もちろん女房もそれがいいと言ってくれましたしな」いま、この国は平和で金持ちで、大学に行こうとしてそこそこ頑張って勉強すれば、ほとんどの人が希望をかなえることができる。こんなことはこの国の歴史の中でかつてなかったと思う。勉強ができて大学に行きたくても、いろいろな事情で行けない人が大量にいたのが、つい三〇年前の状況であった。国外をみてもおそらくこんな国はない。勉強は人によっては厳しく苦しいものである。しかし、一方で内外の状況や過去を思うとき、今持っているすべての時間やエネルギーを勉強に注ぎ込める状況に感謝しないわけにはいかない。

(丹羽健夫著、「眠られぬ受験生のために」、中央公論社)

---