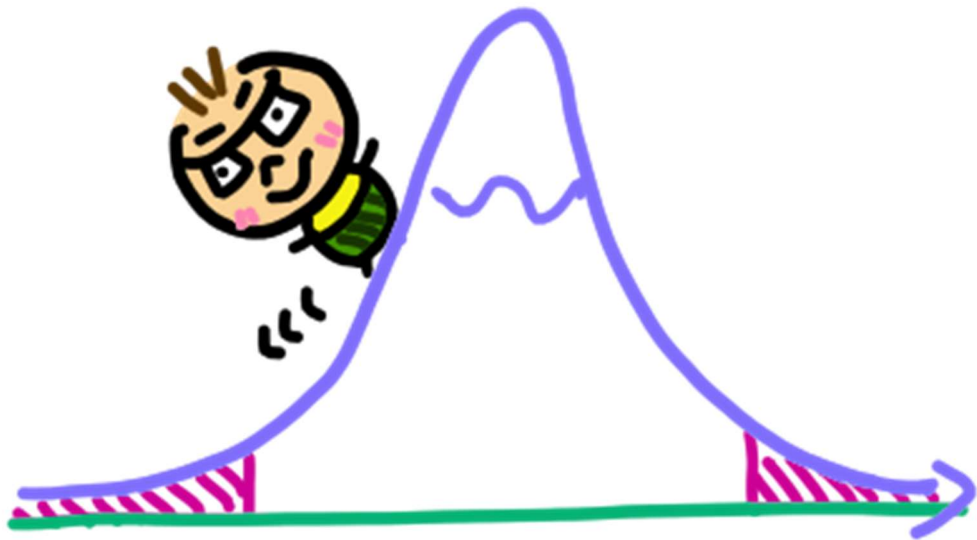


数学 B

15 確率分布と 統計的な推測



< 講義ノート >

確率分布と統計的な推測

1. 確率分布

- (1) 確率変数と確率分布
- (2) 確率変数の期待値と分散
- (3) 確率変数の変換
- (4) 確率変数の和と期待値
- (5) 独立な確率変数
 - (i) 事象の独立と確率変数の独立
 - (ii) 独立な確率変数の積の期待値・和の分散
- (6) 二項分布
- (7) 正規分布
 - (i) 連続型確率変数と確率密度関数
 - (ii) 正規分布と標準正規分布
 - (iii) 二項分布の正規分布による近似

2. 統計的な推測

- (1) 母集団と標本
- (2) 標本平均とその分布
 - (i) 標本平均・標本比率と正規分布
 - (ii) 大数の法則
- (3) 推定
 - (i) 母平均の推定
 - (ii) 母比率の推定
- (4) 仮説検定

1. 確率分布

(1) 確率変数と確率分布

(例) 100円硬貨と10円硬貨を1枚ずつ投げるとき、
表の出た硬貨の金額の和を X 円とする。

金額の種類は、

	X	P
100円が表、10円も表なら	110円	$\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$
100円が表、10円が裏なら	100円	$\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$
100円が裏、10円が表なら	10円	$\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$
100円が裏、10円も裏なら	0円	$\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

また、

X の値	110	100	10	0	計
確率 P	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	1

まとめ

確率変数：試行の結果によって値が定まる変数

確率分布：確率変数 X のとり値に、その値をとる確率を
(または単に「分布」) 対応させたもの

例1 赤球3個と白球5個が入っている袋から、同時に2個の球を取り出すとき、その中に含まれている赤球の個数 X の確率分布を求めよ。

解) X は 0, 1, 2 をとる確率変数である。

$X=0$ は「2個とも白球」

$X=1$ は「赤球と白球が1個ずつ」

$X=2$ は「2個とも赤球」

したがって、それぞれの確率は

$$P(X=0) = \frac{{}_5C_2}{{}_8C_2} = \frac{10}{28}$$

$$P(X=1) = \frac{{}_3C_1 \times {}_5C_1}{{}_8C_2} = \frac{15}{28}$$

$$P(X=2) = \frac{{}_3C_2}{{}_8C_2} = \frac{3}{28}$$

よって、 X の確率分布は以下のようになる。

X	0	1	2	計
P	$\frac{10}{28}$	$\frac{15}{28}$	$\frac{3}{28}$	1

例2

2個のさいころを同時に投げるとき、出る目の和 X についての確率 $P(X=5)$, $P(3 \leq X \leq 5)$ を、それぞれ求めよ。

解) X の確率分布は次の表のようになる。

B ^A	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

X	2	3	4	5	6	7	8	9	10	11	12	合計
P	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	1

$$\text{よって } P(X=5) = \frac{4}{36} = \frac{1}{9} //$$

$$\begin{aligned} P(3 \leq X \leq 5) &= P(X=3) + P(X=4) + P(X=5) \\ &= \frac{2}{36} + \frac{3}{36} + \frac{4}{36} \\ &= \frac{9}{36} = \frac{1}{4} // \end{aligned}$$

(2) 確率変数の期待値と分散 + 標準偏差

<期待値>

確率変数 X 以下の表に示した分布に従うとき、

X	x_1	x_2	x_3		x_n	計
p	p_1	p_2	p_3		p_n	1

$$E(X) = x_1 p_1 + x_2 p_2 + \cdots + x_n p_n = \sum_{k=1}^n x_k p_k$$

を X の「期待値」(または「平均」)という。

例1 2枚の硬貨を同時に投げ、表が出る硬貨の枚数を X とするとき、 X^2 の期待値を求めよ。

解)

$X=0$ となる(2枚とも裏)確率は $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

$X=1$ となる(表1枚裏1枚)確率は $2C_1 \times \frac{1}{2} \times \frac{1}{2} = \frac{2}{4}$

$X=2$ となる(2枚とも表)確率は $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

よって X の確率分布は次のようになる。

X	0	1	2	計
p	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$	1

X^2 の期待値は

$$E(X^2) = 0^2 \times \frac{1}{4} + 1^2 \times \frac{2}{4} + 2^2 \times \frac{1}{4} = \frac{3}{2} //$$

例2 100円硬貨1枚, 50円硬貨2枚を投げ、
表が出た硬貨の金額の和を X とするとき、
 X の期待値を求めよ。

解) X のとりうる値は 0, 50, 100, 150, 200

$$P(X=0) = \left(\frac{1}{2}\right)^3 = \frac{1}{8}$$

$$P(X=50) = \frac{1}{2} \times \left({}_2C_1 \times \frac{1}{2} \times \frac{1}{2}\right) = \frac{2}{8}$$

$$P(X=100) = \frac{1}{2} \times \left(\frac{1}{2}\right)^2 + \frac{1}{2} \times \left(\frac{1}{2}\right)^2 = \frac{2}{8}$$

$$P(X=150) = \frac{1}{2} \times \left({}_2C_1 \times \frac{1}{2} \times \frac{1}{2}\right) = \frac{2}{8}$$

$$P(X=200) = \left(\frac{1}{2}\right)^3 = \frac{1}{8}$$

よって X の確率分布は

右のようになる。

X	0	50	100	150	200	計
P	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{2}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	1

X の期待値は

$$\begin{aligned} E(X) &= 0 \times \frac{1}{8} + 50 \times \frac{2}{8} + 100 \times \frac{2}{8} + 150 \times \frac{2}{8} + 200 \times \frac{1}{8} \\ &= \frac{1}{8} (100 + 200 + 300 + 200) = 100 // \end{aligned}$$

〈分散〉と標準偏差

(復習) n 個のデータ x_i ($i=1, 2, \dots, n$) について,

$$\text{分散} \quad S^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n}$$

$$\text{標準偏差} \quad S = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n}}$$

(まとめ) 確率変数 X の分散・標準偏差 ($E(X) = m$)

$$\text{分散} \quad V(X) = E((X - m)^2)$$

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

$$= \sum_{k=1}^n (x_k - m)^2 p_k$$

$$\text{標準偏差} \quad \sigma(X) = \sqrt{V(X)}$$

$$V(X) = \sum_{k=1}^n (x_k - m)^2 p_k$$

$$= \sum_{k=1}^n (x_k^2 - 2mx_k + m^2) p_k$$

$$= \sum_{k=1}^n x_k^2 p_k - 2m \sum_{k=1}^n x_k p_k + m^2 \sum_{k=1}^n p_k$$

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

例3 2枚の硬貨を同時に投げるとき、表の出る枚数 X の分散 $V(X)$ と標準偏差 $\sigma(X)$ を求めよ。

解) X の確率分布は右の表のようになる。
よってその平均は

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

$$E(X) = 0 \times \frac{1}{4} + 1 \times \frac{2}{4} + 2 \times \frac{1}{4} = 1$$

したがって、 X の分散 $V(X)$ と標準偏差 $\sigma(X)$ は、

$$V(X) = (0-1)^2 \times \frac{1}{4} + (1-1)^2 \times \frac{2}{4} + (2-1)^2 \times \frac{1}{4} = \frac{1}{2} //$$

$$\sigma(X) = \sqrt{\frac{1}{2}} = \frac{\sqrt{2}}{2} //$$

例4 1個のサイコロを投げて出た目の値を X とする。
 X の期待値 $E(X)$ 、分散 $V(X)$ 、標準偏差 $\sigma(X)$ を求めよ。

解) X の確率分布は下の表のようになる。

X	1	2	3	4	5	6	計
P	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	1

$$\text{よって、} E(X) = 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} = \frac{7}{2} //$$

$$\begin{aligned} \text{また、} V(X) &= \frac{1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2}{6} - \left(\frac{7}{2}\right)^2 \\ &= \frac{91}{6} - \frac{49}{4} = \frac{35}{12} // \end{aligned}$$

$$\sigma(X) = \sqrt{\frac{35}{12}} = \frac{\sqrt{105}}{6} //$$

(3) 確率変数の変換

(復習) 変量の変換

a, b は定数とする。変量 x のデータから
 $z = ax + b$ によって新しい変量 z のデータが
得られるとき、 x, z のデータの平均値を $\bar{x}, \bar{z}$,
分散を S_x^2, S_z^2 とすると、

$$\bar{z} = a\bar{x} + b, \quad S_z^2 = a^2 S_x^2, \quad S_z = |a| S_x$$

確率変数 X が右の表に示された
分布に従うとする。

X	x_1	x_2	$\dots$	x_n	計
P	p_1	p_2	$\dots$	p_n	1

a, b が定数のとき、 X の1次式

$$Y = aX + b$$

Y	y_1	y_2	$\dots$	y_n	計
P	p_1	p_2	$\dots$	p_n	1

で Y を定めるとき、 Y も確率変数となる。

このような Y を考えることを「確率変数の変換」という。

< Y の期待値 >

(まとめ) $Y = aX + b$ のとき
 $E(Y) = aE(X) + b$

$$E(Y) = E(aX + b)$$

$$= (ax_1 + b) \times p_1 + (ax_2 + b) \times p_2 + \dots + (ax_n + b) \times p_n$$

$$= ax_1 p_1 + b p_1 + ax_2 p_2 + b p_2 + \dots + ax_n p_n + b p_n$$

$$= a(x_1 p_1 + x_2 p_2 + \dots + x_n p_n) + b(p_1 + p_2 + \dots + p_n)$$

$$= aE(X) + b$$

<Yの分散>

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

まとめ $Y=aX+b$ のとき

$$\begin{aligned} V(Y) &= a^2 V(X) \\ \sigma(Y) &= |a| \sigma(X) \end{aligned}$$

<Yの標準偏差>

$$\begin{aligned} \sigma(Y) &= \sqrt{V(Y)} \\ &= \sqrt{a^2 V(X)} \\ &= |a| \sqrt{V(X)} \end{aligned}$$

まとめ $Y=aX+b$ のとき

$$\begin{cases} E(Y) = aE(X) + b \\ V(Y) = a^2 V(X) \\ \sigma(Y) = |a| \sigma(X) \end{cases}$$

例) 確率変数 X の確率分布が

右の表のように与えられたとき、

確率変数 $Y = -2X + 1$ の

期待値と分散・標準偏差を求めよ。

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

解) 確率変数 X の期待値 $E(X)$ は

$$\begin{aligned} E(X) &= 0 \times \frac{1}{6} + 1 \times \frac{2}{6} + 2 \times \frac{2}{6} + 3 \times \frac{1}{6} \\ &= \frac{2+4+3}{6} = \frac{3}{2} \end{aligned}$$

確率変数 X の分散 $V(X)$ は

$$\begin{aligned} V(X) &= E(X^2) - \{E(X)\}^2 \\ &= 0^2 \times \frac{1}{6} + 1^2 \times \frac{2}{6} + 2^2 \times \frac{2}{6} + 3^2 \times \frac{1}{6} - \left(\frac{3}{2}\right)^2 \\ &= \frac{2+8+9}{6} - \frac{9}{4} = \frac{38-27}{12} \\ &= \frac{19}{6} - \frac{9}{4} = \frac{11}{12} \end{aligned}$$

確率変数 X の標準偏差 $\sigma(X)$ は

$$\begin{aligned} \sigma(X) &= \sqrt{V(X)} \\ &= \sqrt{\frac{11}{12}} \\ &= \frac{\sqrt{11}}{2\sqrt{3}} \\ &= \frac{\sqrt{33}}{6} \end{aligned}$$

まとめ直すと、 $E(X) = \frac{3}{2}$, $V(X) = \frac{11}{12}$, $\sigma(X) = \frac{\sqrt{33}}{6}$

したがって、確率変数 Y の期待値 $E(Y)$ は

$$\begin{aligned} E(Y) &= E(-2X+1) \\ &= -2E(X) + 1 \\ &= -2 \times \frac{3}{2} + 1 = -2 // \end{aligned}$$

確率変数 Y の分散 $V(Y)$ は

$$\begin{aligned} V(Y) &= V(-2X+1) \\ &= (-2)^2 V(X) \\ &= 4 \times \frac{11}{12} = \frac{11}{3} // \end{aligned}$$

確率変数 Y の標準偏差 $\sigma(Y)$ は

$$\begin{aligned} \sigma(Y) &= \sigma(-2X+1) \\ &= |-2| \sigma(X) \\ &= 2 \times \frac{\sqrt{33}}{6} = \frac{\sqrt{33}}{3} // \end{aligned}$$

(4) 確率変数の和と期待値

準備 (同時分布)

2つの確率変数 X, Y について.

X のとる値が $x_1, x_2, \dots, x_n$

Y のとる値が $y_1, y_2, \dots, y_m$

であり,

$$P(X=x_i, Y=y_j) = p_{ij}$$

とおくと、右の表のように

すべての i と j の組み合わせについて

(x_i, y_j) と p_{ij} の対応が得られる。

この対応を X と Y の「同時分布」という。

$X \backslash Y$	y_1	y_2	y_3	$\dots$	$\dots$	y_m	計
x_1	p_{11}	p_{12}	p_{13}	$\dots$	$\dots$	p_{1m}	$p_{1\cdot}$
x_2	p_{21}	p_{22}	p_{23}	$\dots$	$\dots$	p_{2m}	$p_{2\cdot}$
x_3	p_{31}	p_{32}	p_{33}	$\dots$	$\dots$	p_{3m}	$p_{3\cdot}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_n	p_{n1}	p_{n2}	p_{n3}	$\dots$	$\dots$	p_{nm}	$p_{n\cdot}$
計	q_1	q_2	q_3	$\dots$	$\dots$	q_m	1

上の表において,

$$P(X=x_i) = p_{i1} + p_{i2} + p_{i3} + \dots + p_{im} = p_{i\cdot}$$

$$P(Y=y_j) = p_{1j} + p_{2j} + p_{3j} + \dots + p_{nj} = q_j$$

とわかるから、 X と Y はそれぞれ、下の表の分布に従う。

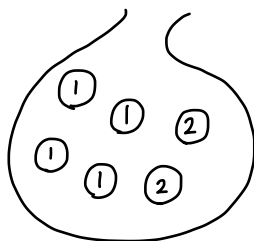
X	x_1	x_2	x_3	$\dots$	x_n	計
P	p_1	p_2	p_3	$\dots$	p_n	1

Y	y_1	y_2	y_3	$\dots$	y_n	計
P	q_1	q_2	q_3	$\dots$	q_n	1



(例) 袋の中に1という番号をつけた4個の玉と2という番号をつけた2個の玉が入っている。この中から最初に1個取り出し、その玉の番号を X とする。取り出した玉を戻さずに2個目を取り出し、その玉の番号を Y とする。 X と Y の同時分布を求めよ。

解) X, Y を確率変数とすると、
 $X=a, Y=b$ となる確率を
 $P(X=a, Y=b)$ と表す。このとき、



$$P(X=1, Y=1) = \frac{4}{6} \times \frac{3}{5} = \frac{12}{30} = \frac{6}{15}$$

$$P(X=1, Y=2) = \frac{4}{6} \times \frac{2}{5} = \frac{8}{30} = \frac{4}{15}$$

$$P(X=2, Y=1) = \frac{2}{6} \times \frac{4}{5} = \frac{8}{30} = \frac{4}{15}$$

$$P(X=2, Y=2) = \frac{2}{6} \times \frac{1}{5} = \frac{2}{30} = \frac{1}{15}$$

よって X と Y の同時分布は右のようになる。

$X \backslash Y$	1	2	計
1	$\frac{6}{15}$	$\frac{4}{15}$	$\frac{10}{15}$
2	$\frac{4}{15}$	$\frac{1}{15}$	$\frac{5}{15}$
計	$\frac{10}{15}$	$\frac{5}{15}$	1

X と Y の確率分布は

X	1	2	計
P	$\frac{10}{15}$	$\frac{5}{15}$	1

Y	1	2	計
P	$\frac{10}{15}$	$\frac{5}{15}$	1

2つの確率変数 X, Y の和を $Z = X + Y$ とするとき、
 Z の期待値 $E(Z)$ について。

例えば、2つの確率変数 X, Y の確率分布が

X	x_1	x_2	計
P	P_{11}	P_{12}	1

Y	y_1	y_2	計
P	q_1	q_2	1

と与えられているとき、右の表のように

$$P(X = x_i, Y = y_j) = P_{ij}$$

とすると、 $P_{11} + P_{12} = P_1$, $P_{21} + P_{22} = P_2$

$P_{11} + P_{21} = q_1$, $P_{12} + P_{22} = q_2$ が成り立つ。

$X = x_i, Y = y_j$ のとき、 $Z = x_i + y_j$ であり、

その確率は P_{ij} だから、

$X \backslash Y$	y_1	y_2	計
x_1	P_{11}	P_{12}	P_1
x_2	P_{21}	P_{22}	P_2
計	q_1	q_2	1

$$E(Z) = (x_1 + y_1) \times P_{11} + (x_1 + y_2) \times P_{12} + (x_2 + y_1) \times P_{21} + (x_2 + y_2) \times P_{22}$$

$$= x_1(P_{11} + P_{12}) + x_2(P_{21} + P_{22}) + y_1(P_{11} + P_{21}) + y_2(P_{12} + P_{22})$$

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

また、

$$E(aX + bY) = E(aX) + E(bY)$$

$$= aE(X) + bE(Y)$$

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

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

まとめ

2つの確率変数 X, Y について

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

a, b を定数としたとき

$$E(aX+bY) = aE(X) + bE(Y)$$

さらに、3つの確率変数 X, Y, Z について

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

(5) 独立な確率変数

(i) 事象の独立と確率変数の独立

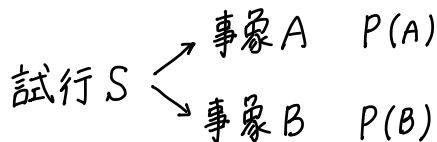
<事象の独立>

(復習) 独立な試行の確率 場合の数と確率

2つの試行 S と T が独立であるとき、

S で事象 A が起こり、かつ T で事象 B が起こる確率 $P(A \cap B)$ は

$$P(A \cap B) = P(A) \times P(B)$$



確率の乗法定理 $P(A \cap B) = P(A) \times P(B)$

$$P(A \cap B) = P(A) \times P(B)$$

まとめ 2つの事象 A, B について $P_A(B) = P(B)$ が成り立つとき、

事象 A, B は独立であるという。

$$\Leftrightarrow P(A \cap B) = P(A) \times P(B)$$

2つの事象 A, B について $P_A(B) \neq P(B)$ が成り立たないとき、

事象 A, B は従属であるという。

例

さいころを1回振る。事象Aと事象Bが次のようなとき、

AとBが互いに独立か従属かを調べよ。

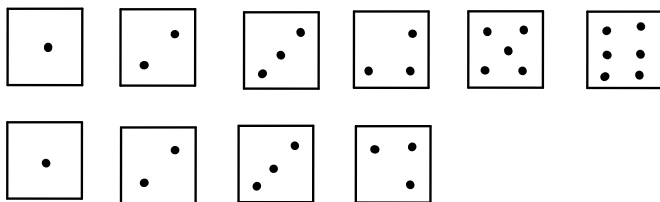
(1) 事象A: 「4以下の目が出る」 事象B: 「偶数の目が出る」

(2) 事象A: 「5以下の目が出る」 事象B: 「偶数の目が出る」

解) (1) $P(A) = \frac{4}{6} = \frac{2}{3}$, $P(B) = \frac{3}{6} = \frac{1}{2}$, $P(A \cap B) = \frac{2}{6} = \frac{1}{3}$

こゝで、 $P(A) \times P(B) = \frac{2}{3} \times \frac{1}{2} = \frac{1}{3} = P(A \cap B)$

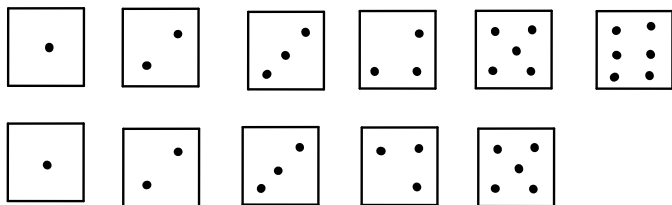
よって、 $P(A \cap B) = P(A) \times P(B)$ が成り立つので、AとBは独立 //



(2) $P(A) = \frac{5}{6}$, $P(B) = \frac{3}{6} = \frac{1}{2}$, $P(A \cap B) = \frac{2}{6} = \frac{1}{3}$

こゝで、 $P(A) \times P(B) = \frac{5}{6} \times \frac{1}{2} = \frac{5}{12} \neq P(A \cap B)$

よって、 $P(A \cap B) \neq P(A) \times P(B)$ となるので、AとBは従属 //



<確率変数の独立>

2つの確率変数 X, Y があり、

X のとり任意の値 a と Y のとり任意の値 b について、

$$P(X=a, Y=b) = P(X=a) \times P(Y=b)$$

が成り立つとき、確率変数 X と Y は互いに独立であるという。

同時分布でみてみよう!

2つの確率変数 X, Y の確率分布が

X	x_1	x_2	計
P	p_1	p_2	1

Y	y_1	y_2	計
P	q_1	q_2	1

$X \backslash Y$	y_1	y_2	計
x_1	p_{11}	p_{12}	p_1
x_2	p_{21}	p_{22}	p_2
計	q_1	q_2	1

で与えられているとき、右の表のように

$P(X=x_i, Y=y_j) = p_{ij}$ とすると、 X と Y が互いに独立である

とは、 $p_{ij} = P(X=x_i, Y=y_j) = P(X=x_i) \times P(Y=y_j) = p_i q_j$

となること。

○「確率変数 X と Y が独立」のとき

事象 A が起これば 1、起こらなければ 0 の値をとる確率変数を X

事象 B が起これば 1、起こらなければ 0 の値をとる確率変数を Y

とすると、

$$P(X=0, Y=0) = P(X=0) \times P(Y=0)$$

$$P(X=1, Y=0) = P(X=1) \times P(Y=0)$$

$$P(X=0, Y=1) = P(X=0) \times P(Y=1)$$

$$P(X=1, Y=1) = P(X=1) \times P(Y=1)$$

$$P(X=1, Y=1) = P(X=1) \times P(Y=1) \text{ は、}$$

$$P(A \cap B) = P(A) \times P(B) \text{ と同値}$$

したがって、

「確率変数 X と Y が独立」ならば「事象 A と B は独立」

○「事象 A と B が独立」のとき

X は 0 と 1 の値しかとらないので、

$$P(X=0) + P(X=1) = 1 \quad \cdots \textcircled{1}$$

$$P(Y=0) + P(Y=1) = 1 \quad \cdots \textcircled{2} \text{ であり、}$$

$$P(X=1) = P(X=1, Y=0) + P(X=1, Y=1) \quad \cdots \textcircled{3}$$

$$P(Y=1) = P(X=0, Y=1) + P(X=1, Y=1) \quad \cdots \textcircled{4}$$

$$\text{また、} P(X=1, Y=1) = P(X=1) \times P(Y=1) \quad \cdots \textcircled{5} \text{ である。}$$

$$\text{他、} P(X=0, Y=0) = P(X=0) \times P(Y=0) \quad \cdots \textcircled{6}$$

$$P(X=1, Y=0) = P(X=1) \times P(Y=0) \quad \cdots \textcircled{7}$$

$$P(X=0, Y=1) = P(X=0) \times P(Y=1) \quad \cdots \textcircled{8}$$

を示したい。

$$P(X=0) + P(X=1) = 1 \quad \cdots \textcircled{1}$$

$$P(Y=0) + P(Y=1) = 1 \quad \cdots \textcircled{2}$$

$$P(X=1) = P(X=1, Y=0) + P(X=1, Y=1) \quad \cdots \textcircled{3}$$

$$P(Y=1) = P(X=0, Y=1) + P(X=1, Y=1) \quad \cdots \textcircled{4}$$

$$P(X=1, Y=1) = P(X=1) \times P(Y=1) \quad \cdots \textcircled{5}$$

$$P(X=0, Y=1) = P(Y=1) - P(X=1, Y=1) \quad (\textcircled{4} \text{より})$$

$$= P(Y=1) - P(X=1) \times P(Y=1) \quad (\textcircled{5} \text{より})$$

$$= P(Y=1) \{1 - P(X=1)\} = P(X=0) \times P(Y=1) \quad \cdots \textcircled{6}$$

$$P(X=1, Y=0) = P(X=1) - P(X=1, Y=1) \quad (\textcircled{3} \text{より})$$

$$= P(X=1) - P(X=1) \times P(Y=1) \quad (\textcircled{5} \text{より})$$

$$= P(X=1) \{1 - P(Y=1)\} = P(X=1) \times P(Y=0) \quad \cdots \textcircled{7}$$

$$P(X=0, Y=0) = P(X=0) - P(X=0, Y=1)$$

$$= P(X=0) - P(X=0) \times P(Y=1) \quad (\textcircled{6} \text{より})$$

$$= P(X=0) \{1 - P(Y=1)\} = P(X=0) \times P(Y=0) \quad \cdots \textcircled{8}$$

よって、 $P(X=a, Y=b) = P(X=a) \times P(Y=b)$ となり、

「事象 A と B が独立」ならば「確率変数 X と Y は独立」

とわかる。

事象 A と B が独立であることと、

対応する確率変数 X と Y が独立であること

は同値である。

(5) 独立な確率変数

(ii) 独立な確率変数の積の期待値・和の分散

(復習) 確率変数の独立

2つの確率変数 X, Y があり、

X のとる任意の値 a と Y のとる任意の値 b について、

$$P(X=a, Y=b) = P(X=a) \times P(Y=b)$$

が成り立つとき、

確率変数 X と Y は 互いに独立であるという。

同時分布で見てみよう!

2つの確率変数 X, Y の確率分布が

X	x_1	x_2	計
P	p_1	p_2	1

Y	y_1	y_2	計
P	q_1	q_2	1

$X \backslash Y$	y_1	y_2	計
x_1	p_{11}	p_{12}	p_1
x_2	p_{21}	p_{22}	p_2
計	q_1	q_2	1

で与えられているとき、右の表のように

$$P(X=x_i, Y=y_j) = p_{ij} \text{ とすると、}$$

X と Y が 互いに独立であるとは、

$$p_{ij} = P(X=x_i, Y=y_j) = P(X=x_i) \times P(Y=y_j) = p_i q_j$$

となること。

まとめ1

確率変数 X と Y が互いに独立ならば

$$E(XY) = E(X)E(Y)$$

$$V(X+Y) = V(X) + V(Y)$$

証明

確率変数 X と Y が互いに独立であるとき、

X と Y の同時分布は

$X \backslash Y$	y_1	y_2	y_3	$\dots$	y_m	計
x_1	$p_1 q_1$	$p_1 q_2$	$p_1 q_3$	$\dots$	$p_1 q_m$	p_1
x_2	$p_2 q_1$	$p_2 q_2$	$p_2 q_3$	$\dots$	$p_2 q_m$	p_2
x_3	$p_3 q_1$	$p_3 q_2$	$p_3 q_3$	$\dots$	$p_3 q_m$	p_3
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
x_n	$p_n q_1$	$p_n q_2$	$p_n q_3$	$\dots$	$p_n q_m$	p_n
計	q_1	q_2	q_3	$\dots$	q_m	1

$$\begin{aligned} E(XY) &= x_1 y_1 \times p_1 q_1 + x_1 y_2 \times p_1 q_2 + \dots + x_1 y_m \times p_1 q_m \\ &\quad + x_2 y_1 \times p_2 q_1 + x_2 y_2 \times p_2 q_2 + \dots + x_2 y_m \times p_2 q_m \\ &\quad + x_3 y_1 \times p_3 q_1 + x_3 y_2 \times p_3 q_2 + \dots + x_3 y_m \times p_3 q_m \\ &\quad + \dots \\ &\quad \vdots \\ &\quad + x_n y_1 \times p_n q_1 + x_n y_2 \times p_n q_2 + \dots + x_n y_m \times p_n q_m \\ &= x_1 p_1 (y_1 q_1 + y_2 q_2 + \dots + y_m q_m) \\ &\quad + x_2 p_2 (y_1 q_1 + y_2 q_2 + \dots + y_m q_m) \\ &\quad + x_3 p_3 (y_1 q_1 + y_2 q_2 + \dots + y_m q_m) \end{aligned}$$

$$\begin{aligned}
& + x_n p_n (y_1 q_1 + y_2 q_2 + \dots + y_m q_m) \\
& = (x_1 p_1 + x_2 p_2 + \dots + x_n p_n) (y_1 q_1 + y_2 q_2 + \dots + y_m q_m) \\
& = E(X) \times E(Y)
\end{aligned}$$

$$\begin{aligned}
V(X+Y) &= E((X+Y)^2) - \{E(X+Y)\}^2 \\
&= E(X^2 + 2XY + Y^2) - \{E(X) + E(Y)\}^2 \\
&= E(X^2) + 2E(XY) + E(Y^2) - \{E(X)\}^2 - 2E(X)E(Y) - \{E(Y)\}^2 \\
&= E(X^2) - \{E(X)\}^2 + E(Y^2) - \{E(Y)\}^2 = V(X) + V(Y)
\end{aligned}$$

例2

確率変数 X, Y, Z が互いに独立ならば

$$E(XYZ) = E(X)E(Y)E(Z)$$

$$V(X+Y+Z) = V(X) + V(Y) + V(Z)$$

さらに、

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

例1 大小2個のさいころを投げ、出る目をそれぞれ X, Y とする。

(1) XY の期待値を求めよ。

(2) $X+Y$ および $2X+4Y$ の分散を求めよ。

解)

(1) 確率変数 X と Y は独立であるから、

$$E(XY) = E(X) \times E(Y)$$

$$\therefore E(X) = E(Y)$$

$$= 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \frac{7}{2}$$

$$\therefore E(XY) = \frac{7}{2} \times \frac{7}{2} = \frac{49}{4} //$$

$$(2) V(X+Y) = V(X) + V(Y)$$

$$= 2 \times \left\{ \frac{1}{6} (1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) - \left(\frac{7}{2}\right)^2 \right\}$$

$$= 2 \times \left(\frac{91}{6} - \frac{49}{4} \right) = 2 \times \frac{35}{12} = \frac{35}{6} //$$

$$V(2X+4Y) = 2^2 V(X) + 4^2 V(Y)$$

$$= 4 \times \frac{35}{12} + 16 \times \frac{35}{12} = \frac{175}{3} //$$

例2 確率分布が下の表のような3つの確率変数 X, Y, Z が互いに独立であるとき、 XYZ の期待値および、 $X+Y+Z$ の分散を求めよ。

X	0	1	合計
P	$\frac{1}{2}$	$\frac{1}{2}$	1

Y	0	1	合計
P	$\frac{2}{3}$	$\frac{1}{3}$	1

Z	0	1	合計
P	$\frac{5}{6}$	$\frac{1}{6}$	1

$$\text{解) } E(X) = 0 \times \frac{1}{2} + 1 \times \frac{1}{2} = \frac{1}{2}$$

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

$$E(Z) = 0 \times \frac{5}{6} + 1 \times \frac{1}{6} = \frac{1}{6} \quad \text{よ、}$$

$$E(XYZ) = E(X) \times E(Y) \times E(Z)$$

$$= \frac{1}{2} \times \frac{1}{3} \times \frac{1}{6}$$

$$= \frac{1}{36} //$$

次に、

$$V(X) = E(X^2) - \{E(X)\}^2 = \left(0^2 \times \frac{1}{2} + 1^2 \times \frac{1}{2}\right) - \left(\frac{1}{2}\right)^2 = \frac{1}{4}$$

$$V(Y) = E(Y^2) - \{E(Y)\}^2 = \left(0^2 \times \frac{2}{3} + 1^2 \times \frac{1}{3}\right) - \left(\frac{1}{3}\right)^2 = \frac{2}{9}$$

$$V(Z) = E(Z^2) - \{E(Z)\}^2 = \left(0^2 \times \frac{5}{6} + 1^2 \times \frac{1}{6}\right) - \left(\frac{1}{6}\right)^2 = \frac{5}{36}$$

$$\text{よ、} V(X+Y+Z) = V(X) + V(Y) + V(Z)$$

$$= \frac{1}{4} + \frac{2}{9} + \frac{5}{36} = \frac{11}{18} //$$

(6) 二項分布

まずは基本から

例) 1つのサイコロを3回投げたとき、2の目が出る回数を X とすると、

2の目が r 回 ($r=0,1,2,3$) 出る確率 $P(X=r)$ は

$$P(X=r) = {}_3C_r \left(\frac{1}{6}\right)^r \left(\frac{5}{6}\right)^{3-r}$$

よって X の確率分布は

X	0	1	2	3	計
P	${}_3C_0 \left(\frac{5}{6}\right)^3$	${}_3C_1 \times \frac{1}{6} \times \left(\frac{5}{6}\right)^2$	${}_3C_2 \times \left(\frac{1}{6}\right)^2 \times \frac{5}{6}$	${}_3C_3 \left(\frac{1}{6}\right)^3$	1

二項定理

$${}_nC_0 a^n + {}_nC_1 a^{n-1} b^1 + {}_nC_2 a^{n-2} b^2 + \dots + {}_nC_{n-1} a b^{n-1} + {}_nC_n b^n = (a+b)^n$$

まとめ

1回の試行で事象 A の起こる確率を P とすると、

事象 A が起こらない確率は $q = 1 - P$ である。

これを n 回行う反復試行において、

事象 A の起こる回数を X とすると、 $X=r$ となる確率 $P(X=r)$ は、

$$P(X=r) = {}_nC_r P^r q^{n-r} \quad (\text{ただし、} q = 1 - p)$$

したがって、確率変数 X の確率分布は次の表のようになる。

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

このような確率分布を二項分布といい、 $B(n, p)$ で表す。

このとき、確率変数 X は二項分布 $B(n, p)$ に従う という。

例1

1枚の硬貨を5回投げるとき、表の出る回数を X とする。

このとき X はどのような二項分布に従うか。

また、確率 $P(1 \leq X \leq 3)$ を求めよ。

解) 硬貨を1回投げるとき、表の出る確率は $\frac{1}{2}$ だから、

X は二項分布 $B(5, \frac{1}{2})$ に従う。

$$\text{また、 } P(X=r) = {}_5C_r \left(\frac{1}{2}\right)^r \left(1-\frac{1}{2}\right)^{5-r}$$

なので、確率変数 X の確率分布は以下のようになる。

X	0	1	2	3	4	5	計
P	${}_5C_0 \left(\frac{1}{2}\right)^5$	${}_5C_1 \times \frac{1}{2} \times \left(\frac{1}{2}\right)^4$	${}_5C_2 \times \left(\frac{1}{2}\right)^2 \times \left(\frac{1}{2}\right)^3$	${}_5C_3 \times \left(\frac{1}{2}\right)^3 \times \left(\frac{1}{2}\right)^2$	${}_5C_4 \times \left(\frac{1}{2}\right)^4 \times \frac{1}{2}$	${}_5C_5 \times \left(\frac{1}{2}\right)^5$	1

したがって、

$$\begin{aligned} P(1 \leq X \leq 3) &= P(X=1) + P(X=2) + P(X=3) \\ &= {}_5C_1 \times \frac{1}{2} \times \left(\frac{1}{2}\right)^4 + {}_5C_2 \times \left(\frac{1}{2}\right)^2 \times \left(\frac{1}{2}\right)^3 + {}_5C_3 \times \left(\frac{1}{2}\right)^3 \times \left(\frac{1}{2}\right)^2 \\ &= 5 \times \left(\frac{1}{2}\right)^5 + 10 \times \left(\frac{1}{2}\right)^5 + 10 \times \left(\frac{1}{2}\right)^5 \\ &= 25 \times \left(\frac{1}{2}\right)^5 = \frac{25}{32} // \end{aligned}$$

次に、二項分布の「平均」と「分散」

確率変数 X が二項分布 $B(n, p)$ に従うとき、

$$E(X) = np, \quad V(X) = npq, \quad \sigma(X) = \sqrt{npq}$$

(説明) 1回の試行で事象Aの起こる確率が p である試行を
 n 回行うとき、第 k 回目の試行で事象Aが起これば1,
起こらなければ0の値をとる確率変数を X_k とする。

このとき、 $P(X_k=1)=p$, $P(X_k=0)=q$ (ただし、 $q=1-p$)

したがって、 $E(X_k)=1 \times p + 0 \times q = p$

$$\begin{aligned}\text{次に、} V(X_k) &= E(X_k^2) - \{E(X_k)\}^2 \\ &= (1^2 \times p + 0^2 \times q) - p^2 \\ &= p - p^2 = p(1-p) = pq\end{aligned}$$

ここで、 $X = X_1 + X_2 + \dots + X_n$ とおくと、確率変数 X は

n 回の反復試行においてAが起こる回数を示すので、

二項分布に従う。よって、

$$\begin{aligned}E(X) &= E(X_1 + X_2 + \dots + X_n) \\ &= E(X_1) + E(X_2) + \dots + E(X_n) = np\end{aligned}$$

$$\begin{aligned}V(X) &= V(X_1 + X_2 + \dots + X_n) \\ &= V(X_1) + V(X_2) + \dots + V(X_n) = npq\end{aligned}$$

標準偏差 $\sigma(X)$ は、 $\sigma(X) = \sqrt{V(X)} = \sqrt{npq}$

例2 1個のさいころを4回投げ、1の目が出る回数を X とする。
次の値を求めよ。

(1) $P(X=2)$

(2) X の期待値 $E(X)$

(3) X の分散 $V(X)$

(4) X の標準偏差 $\sigma(X)$

解)

(1) X は二項分布 $B(4, \frac{1}{6})$ に従う。

X	0	1	2	3	4	計
P			●			1

$$P(X=2) = {}_4C_2 \times \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right)^2 = \frac{25}{216} //$$

$$(2) E(X) = 4 \times \frac{1}{6} = \frac{2}{3} //$$

$$(3) V(X) = 4 \times \frac{1}{6} \times \frac{5}{6} = \frac{5}{9} //$$

$$(4) \sigma(X) = \sqrt{\frac{5}{9}} = \frac{\sqrt{5}}{3} //$$

(例3) 確率変数 X が二項分布 $B(200, \frac{1}{8})$ に従うとき、

確率変数 $Y = -\frac{4X-65}{5}$ の期待値 $E(Y)$ と分散 $V(Y)$ を求めよ。

解)

確率変数 X は二項分布 $B(200, \frac{1}{8})$ に従うの2''.

$$E(X) = 200 \times \frac{1}{8} = 25 //$$

$$V(X) = 200 \times \frac{1}{8} \times \frac{7}{8} = \frac{175}{8} //$$

$$E(Y) = E\left(-\frac{4X-65}{5}\right)$$

$$= -\frac{4}{5} E(X) + 13$$

$$= -\frac{4}{5} \times 25 + 13 = -7 //$$

$$V(Y) = V\left(-\frac{4X-65}{5}\right)$$

$$= \left(-\frac{4}{5}\right)^2 V(X)$$

$$= \frac{16}{25} \times \frac{175}{8} = 14 //$$

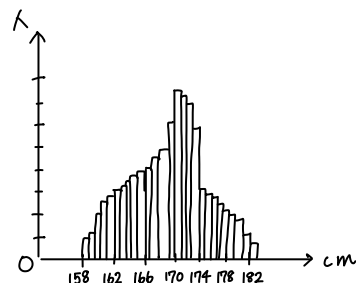
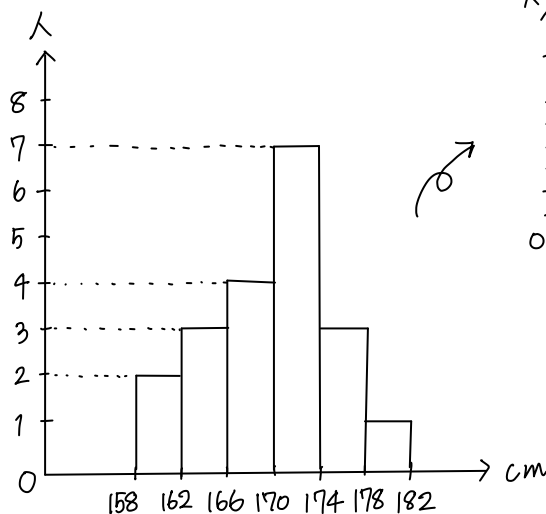
(7) 正規分布

(i) 確率密度関数

準備 (連続型確率変数)

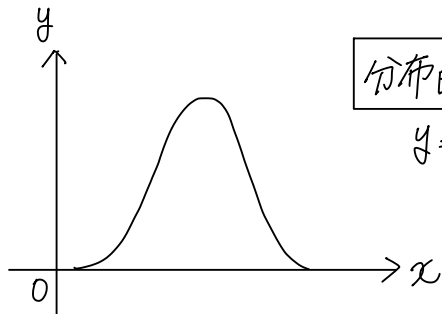
(例)

階級値	階級 (cm)	度数	相対度数
160	158 cm 以上 162 未満	2	0.10
164	162 ~ 166	3	0.15
168	166 ~ 170	4	0.20
172	170 ~ 174	7	0.35
176	174 ~ 178	3	0.15
180	178 ~ 182	1	0.05
	計	20	1.00



離散型
確率変数

データの大きさに → 増 階級の幅 → 狭
↓
曲線に近づく



連続型確率変数

分布曲線

$$y = f(x)$$

確率密度関数



確率密度関数 $y = f(x)$ の性質

[性質1] 常に $f(x) \geq 0$

[性質2] 確率 $P(a \leq x \leq b)$ は、

$y = f(x)$ のグラフと x 軸および 2 直線 $x = a$, $x = b$ で囲まれた部分の面積に等しい。すなわち、

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

[性質3] X のとり値の範囲が $\alpha \leq x \leq \beta$ のとき

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

(例1) 確率変数 X の確率密度関数 $f(x)$ が

$$f(x) = x + \frac{1}{2} \quad (0 \leq x \leq 1)$$

と与えられるとき、次の確率を求めよ。

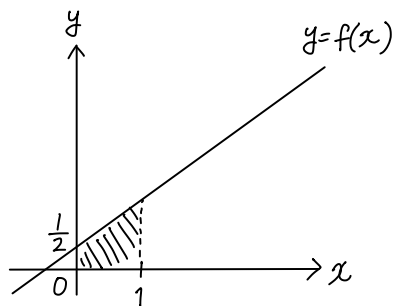
(1) $P(0 \leq X \leq 1)$

(2) $P(0 \leq X \leq 0.5)$

(3) $P(a \leq X \leq b) \quad (0 \leq a < b \leq 1)$

解)

$$\begin{aligned} (1) P(0 \leq X \leq 1) &= \int_0^1 \left(x + \frac{1}{2}\right) dx \\ &= \left[\frac{1}{2}x^2 + \frac{1}{2}x\right]_0^1 \\ &= \frac{1}{2} + \frac{1}{2} = 1 // \end{aligned}$$

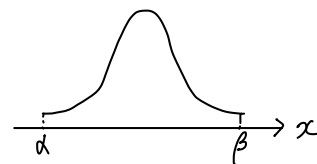


$$\begin{aligned} (2) P(0 \leq X \leq 0.5) &= \int_0^{\frac{1}{2}} \left(x + \frac{1}{2}\right) dx \\ &= \left[\frac{1}{2}x^2 + \frac{1}{2}x\right]_0^{\frac{1}{2}} \\ &= \frac{1}{8} + \frac{1}{4} = \frac{3}{8} = 0.375 // \end{aligned}$$

$$\begin{aligned} (3) P(a \leq X \leq b) &= \int_a^b \left(x + \frac{1}{2}\right) dx \\ &= \left[\frac{1}{2}x^2 + \frac{1}{2}x\right]_a^b \\ &= \frac{1}{2}(b^2 - a^2) + \frac{1}{2}(b - a) \\ &= \frac{1}{2}(b+a)(b-a) + \frac{1}{2}(b-a) \\ &= \frac{1}{2}(b-a)(b+a+1) // \end{aligned}$$

連続型確率変数の「期待値」と「分散」

確率変数 X のとり値が $\alpha \leq x \leq \beta$ で、
その確率密度関数が $f(x)$ のとき、 X の



期待値 $E(X) = \int_{\alpha}^{\beta} x f(x) dx$

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

また、標準偏差 $\sigma(X) = \sqrt{V(X)}$

$V(X)$ について

$$\begin{aligned} V(X) &= \int_{\alpha}^{\beta} (x-m)^2 f(x) dx \\ &= \int_{\alpha}^{\beta} (x^2 - 2mx + m^2) f(x) dx \\ &= \int_{\alpha}^{\beta} x^2 f(x) dx - 2m \int_{\alpha}^{\beta} x f(x) dx + m^2 \int_{\alpha}^{\beta} f(x) dx \\ &= \int_{\alpha}^{\beta} x^2 f(x) dx - m^2 \end{aligned}$$

したがって、

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

例2 確率変数 X の確率密度関数 $f(x)$ が

$$f(x) = \frac{2}{25}x \quad (0 \leq x \leq 5)$$

と与えられるとき、 X の期待値 $E(X)$ と分散 $V(X)$ を求めよ。

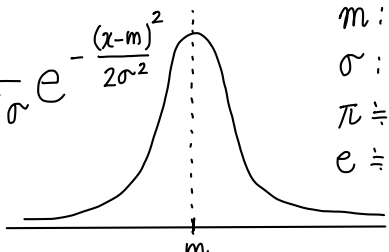
解)

$$\begin{aligned} E(X) &= \int_0^5 x f(x) dx \\ &= \int_0^5 x \times \frac{2}{25} x dx \\ &= \frac{2}{25} \left[\frac{1}{3} x^3 \right]_0^5 \end{aligned} \quad \rightarrow \quad \begin{aligned} &= \frac{2}{25} \times \frac{1}{3} \times 5^3 \\ &= \frac{10}{3} // \end{aligned}$$

$$\begin{aligned} V(X) &= \int_0^5 x^2 f(x) dx - \{E(X)\}^2 \\ &= \int_0^5 x^2 f(x) dx - \left(\frac{10}{3}\right)^2 \\ &= \int_0^5 x^2 \times \frac{2}{25} x dx - \frac{100}{9} \end{aligned} \quad \rightarrow \quad \begin{aligned} &= \frac{2}{25} \left[\frac{1}{4} x^4 \right]_0^5 - \frac{100}{9} \\ &= \frac{2}{25} \times \frac{1}{4} \times 5^4 - \frac{100}{9} \\ &= \frac{25}{2} - \frac{100}{9} = \frac{25}{18} // \end{aligned}$$

(ii) 正規分布と標準正規分布

連続型確率変数の分布の代表例「正規分布」

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}$$


m : 期待値
 σ : 標準偏差
 $\pi \doteq 3.14$
 $e \doteq 2.718$ (ネイピア数)

連続型確率変数 X の確率密度関数 $f(x)$ が m, σ を定数として

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

このとき X は正規分布 $N(m, \sigma^2)$ に従う、という。

$y = f(x)$ を「正規分布曲線」という。

正規分布曲線の性質

[その1] 直線 $x=m$ に関して対称であり、
 $f(x)$ の値は $x=m$ で最大となる。

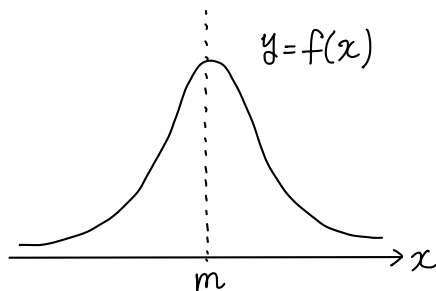
[その2] x 軸を漸近線とする。

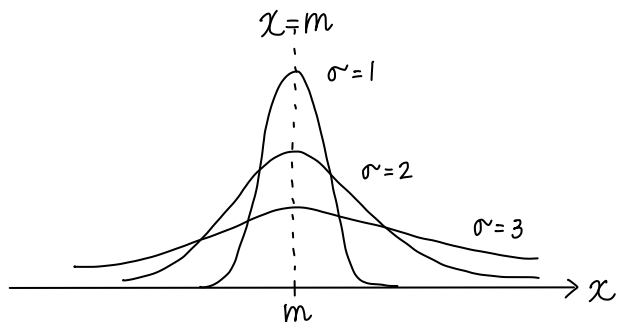
[その3] 標準偏差 σ が

大きくなると、曲線の山が低くたつ横に広がり、

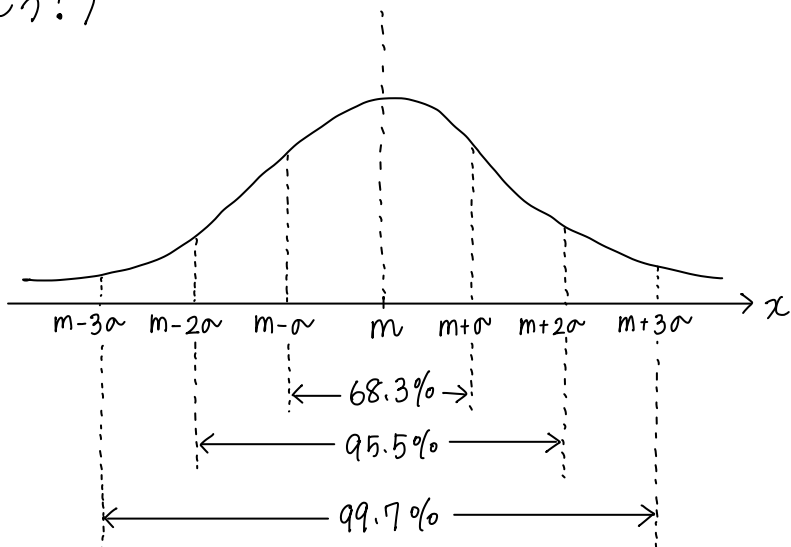
小さくなると、曲線の山は高くなつて対称軸の周りに集まる。

($x=m$)

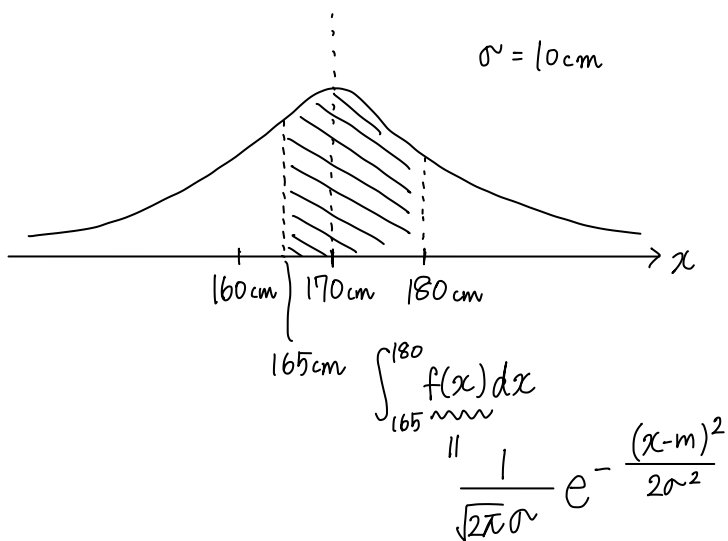




(知っておこう!)



(例)

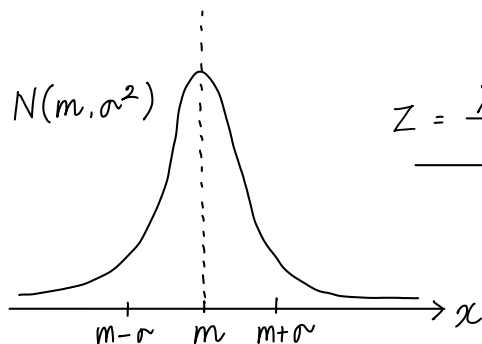


(ii) 正規分布と標準正規分布

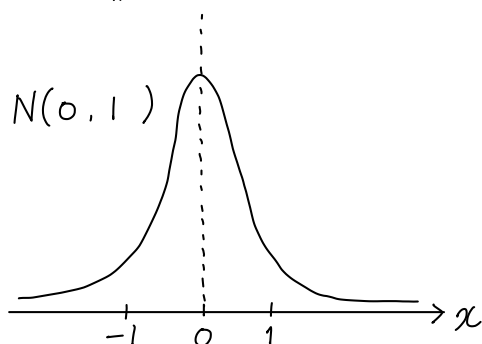
標準化

$$Z = \frac{X - m}{\sigma} = Y$$

標準正規分布

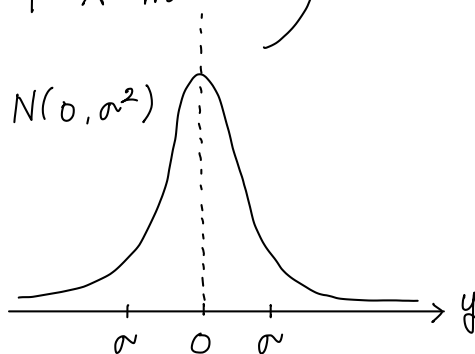


$$Z = \frac{X - m}{\sigma} \longrightarrow$$



$$\searrow Y = X - m$$

$$\nearrow Z = \frac{X - m}{\sigma}$$



$$Z = \frac{X - m}{\sigma} = \frac{1}{\sigma}x + \frac{-m}{\sigma}$$

$$\bar{Z} = \frac{1}{\sigma}\bar{x} + \frac{-m}{\sigma}$$

$$S_Z = \left| \frac{1}{\sigma} \right| S_X = 1$$

$$Z = ax + b \text{ のとき}$$

$$\bar{Z} = a\bar{x} + b$$

$$S_Z = |a| S_X$$

$$f(x) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} \quad \xrightarrow{z = \frac{x-m}{\sigma}} \quad f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

(証明) $a < b$ のとき $P(a \leq z \leq b)$ について考える。

X が $N(m, \sigma^2)$ に従うとき、

$$\begin{aligned} P(a \leq z \leq b) &= P\left(a \leq \frac{X-m}{\sigma} \leq b\right) \\ &= P(m + a\sigma \leq X \leq m + b\sigma) \\ &= \int_{m+a\sigma}^{m+b\sigma} \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} dx \end{aligned}$$

∴ z 、 $z = \frac{x-m}{\sigma}$ とおくと

$$x = m + \sigma z \text{ より } dx = \sigma dz$$

$$\begin{array}{c|c} x & m+a\sigma \rightarrow m+b\sigma \\ \hline z & a \rightarrow b \end{array}$$

$$\begin{aligned} P(a \leq z \leq b) &= \int_a^b \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{1}{2} z^2} \sigma dz \\ &= \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \end{aligned}$$

$$\therefore \text{つまり } f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \text{ となる。} \quad \blacksquare$$

例1 確率変数 X が正規分布 $N(1, 2^2)$ に従うとき、
次の確率を求めよ。

$$(1) P(X \geq 2) \quad (2) P(2 \leq X \leq 3) \quad (3) P(-2 \leq X \leq 2)$$

解) 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。

$Z = \frac{X - 1}{2}$ とおくと、 Z は $N(0, 1)$ に従う。

$$(1) P(X \geq 2) = P(Z \geq \frac{2-1}{2})$$

$$= P(Z \geq 0.5)$$

$$= 0.5 - 0.1915 = 0.3085 //$$

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

$$= P(0.5 \leq Z \leq 1)$$

$$= \Phi(1) - \Phi(0.5) = 0.1498 //$$

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

$$= P(-1.5 \leq Z \leq 0.5)$$

$$= \Phi(0.5) + \Phi(1.5)$$

$$= 0.6247 //$$

例2 ある高校の男子生徒の身長が、平均170.2 cm、標準偏差 5.6 cm の正規分布に従うとする。このとき、身長が 180 cm 以上の生徒は、この高校でおよそ何% いるか。四捨五入して上から2桁の概数で答えよ。

解) 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。

$Z = \frac{X - 170.2}{5.6}$ とおくと、 Z は $N(0, 1)$ に従う。

$$P(X \geq 180) = P\left(Z \geq \frac{180 - 170.2}{5.6}\right)$$

$$= P(Z \geq 1.75)$$

$$= 0.5 - \Phi(1.75) = 0.0401$$

よって、およそ 4.0 % //

(iii) 二項分布の正規分布による近似

(例) 1枚の硬貨を5回投げるとき、表の出る回数を X とする。
このとき X はどのような二項分布に従うか。

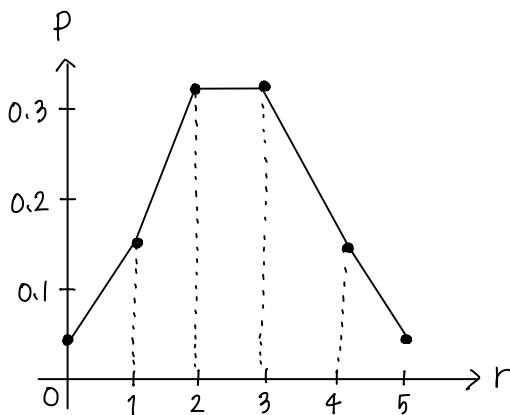
解) 硬貨を1回投げるとき、表の出る確率は $\frac{1}{2}$ だから、

X は二項分布 $B(5, \frac{1}{2})$ に従う。

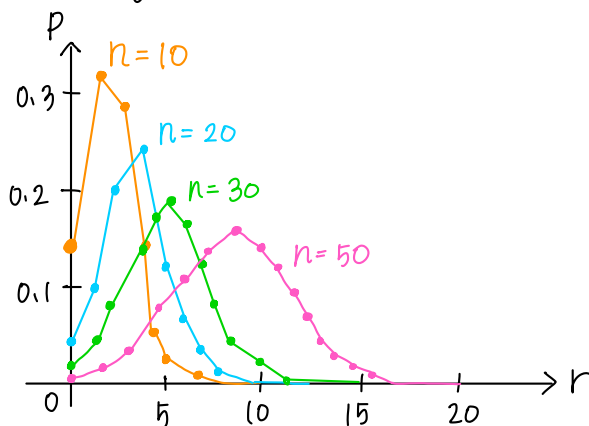
$$\text{また、} P(X=r) = {}_5C_r \left(\frac{1}{2}\right)^r \left(1 - \frac{1}{2}\right)^{5-r}$$

よって、確率変数 X の確率分布は以下になる。

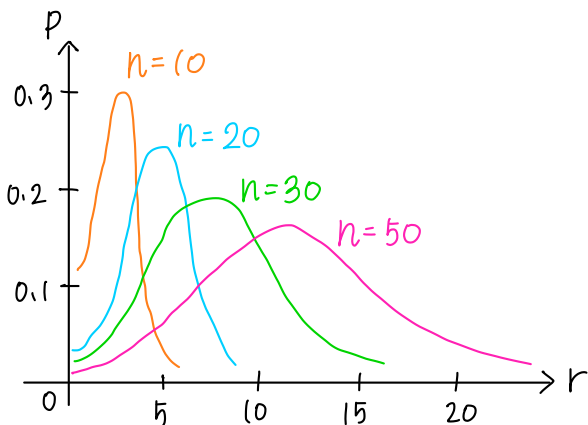
X	0	1	2	3	4	5	計
P	${}_5C_0 \left(\frac{1}{2}\right)^5$	${}_5C_1 \times \frac{1}{2} \times \left(\frac{1}{2}\right)^4$	${}_5C_2 \times \left(\frac{1}{2}\right)^2 \times \left(\frac{1}{2}\right)^3$	${}_5C_3 \times \left(\frac{1}{2}\right)^3 \times \left(\frac{1}{2}\right)^2$	${}_5C_4 \times \left(\frac{1}{2}\right)^4 \times \frac{1}{2}$	${}_5C_5 \times \left(\frac{1}{2}\right)^5$	1



①例 $B(n, \frac{1}{6})$ の場合



$$\begin{cases} E(X) = np \\ V(X) = npq \\ \sigma(X) = \sqrt{npq} \end{cases}$$



まとめ

1. 二項分布 $B(n, p)$ に従う確率変数 X は n が大きいとき、近似的に正規分布 $N(np, npq)$ に従う。
2. 二項分布 $B(n, p)$ に従う確率変数 X に対し、

$$Z = \frac{X - np}{\sqrt{npq}} \text{ は、} n \text{ が大きいとき、近似的に}$$

標準正規分布 $N(0, 1)$ に従う。

(例) 1個のさいころを200回投げるとき、3の倍数の目の出る回数 X が55回以下となる確率を求めよ。

解)

3の倍数の目の出る回数を X とすると、

X は二項分布 $B(200, \frac{1}{3})$ に従う。

その平均を m 、標準偏差を σ とすると、

$$m = 200 \times \frac{1}{3} = \frac{200}{3}$$

$$\sigma = \sqrt{200 \times \frac{1}{3} \times \frac{2}{3}} = \frac{20}{3}$$

$n (= 200)$ は大きいので、

X は正規分布 $N(\frac{200}{3}, (\frac{20}{3})^2)$ に従う。

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。

$$Z = \frac{X - \frac{200}{3}}{\frac{20}{3}} = \frac{3X - 200}{20} \quad \text{とおくと、} Z \text{ は } N(0, 1) \text{ に従う。}$$

$$X \leq 55 \text{ のとき } Z \leq \frac{3 \times 55 - 200}{20} = \frac{-35}{20} = -1.75$$

$$P(X \leq 55) = P(Z \leq -1.75)$$

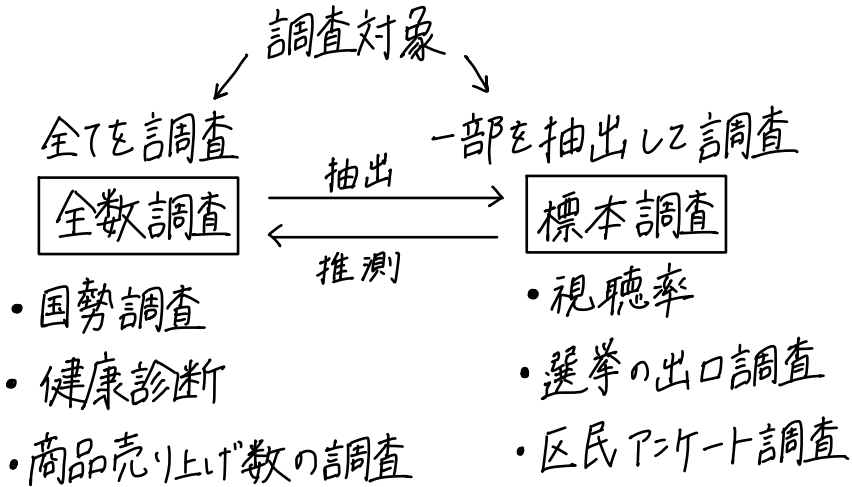
$$= P(Z \geq 1.75)$$

$$= 0.5 - P(0 \leq Z \leq 1.75) = 0.0401 //$$

2. 統計的な推測

(1) 母集団と標本

(i) 統計的な調査



メリット

誤差のない正確な
調査結果が得られる

規模の大きな調査対象
を効率的に調べられる

デメリット

1回の調査に莫大な
時間と労力がかかる

調査結果に大きな誤差
が生ずる可能性がある

標本調査を行うとき、調査の対象全体を「母集団」

母集団から抜き出された要素の集合を「標本」という。

母集団から標本を抜き出すことを、標本の「抽出」という。

また、母集団、標本の要素の個数をもれどれ

「母集団の大きさ」、「標本の大きさ」という。

(ii) 母集団分布

大きき N の母集団において、変量 X の値が、 $x_1, x_2, \dots, x_n$ であり、それぞれの要素の個数を $f_1, f_2, \dots, f_n$ とすると、その度数分布表は右のようになる。

X	度数
x_1	f_1
x_2	f_2
$\vdots$	$\vdots$
x_n	f_n
計	N

この母集団から無作為に1個を取り出したとき、

それが $X = x_i$ となる確率は $P(X = x_i) = \frac{f_i}{N}$ ($i = 1, 2, \dots, n$) であり、 X の確率分布は以下のようなになる。

X	x_1	x_2	x_3	$\dots\dots\dots$	x_n	計
p	$\frac{f_1}{N}$	$\frac{f_2}{N}$	$\frac{f_3}{N}$		$\frac{f_n}{N}$	1

この確率分布を「母集団分布」という。

母集団分布の平均、分散、標準偏差をそれぞれ

「母平均」、「母分散」、「母標準偏差」といい、 m , σ^2 , σ で表す。

例1 1, 2, 3の数字を1つずつ書いた札が、それぞれ

2枚, 5枚, 3枚ある。この10枚の札を母集団とし、

札に書かれた数字 X をこの母集団の変量とするとき、

この母集団の母平均 m , 母分散 σ^2 , 母標準偏差 σ を求めよ。

解)

母集団分布は右のようになる。

X	1	2	3	計
P	$\frac{2}{10}$	$\frac{5}{10}$	$\frac{3}{10}$	1

$$\begin{aligned}\text{よって } m &= 1 \times \frac{2}{10} + 2 \times \frac{5}{10} + 3 \times \frac{3}{10} \\ &= \frac{21}{10} //\end{aligned}$$

$$\begin{aligned}\sigma^2 &= 1^2 \times \frac{2}{10} + 2^2 \times \frac{5}{10} + 3^2 \times \frac{3}{10} - \left(\frac{21}{10}\right)^2 \\ &= \frac{2 + 20 + 27}{10} - \left(\frac{21}{10}\right)^2 \\ &= \frac{490}{100} - \frac{441}{100} = \frac{49}{100} //\end{aligned}$$

$$\sigma = \sqrt{\frac{49}{100}} = \frac{7}{10} //$$

(iii) 「復元抽出」と「非復元抽出」

例2 10本中3本当たりのくじがある。3人が次々1本ずつひくとき、2人だけが当たる確率を次の(1), (2)の状況で考える。

(1) ひいたくじを元に戻すとき

(2) ひいたくじを元に戻さないとき

解)

$$(1) {}_3C_2 \left(\frac{3}{10}\right)^2 \times \frac{7}{10} = \frac{189}{1000} = 0.189$$

$$(2) \begin{array}{ccccc} \text{あ} & \text{あ} & \text{は} & \text{あ} & \text{は} & \text{あ} & \text{は} & \text{あ} & \text{あ} \\ \frac{3}{10} \times \frac{2}{9} \times \frac{7}{8} & + & \frac{3}{10} \times \frac{7}{9} \times \frac{2}{8} & + & \frac{7}{10} \times \frac{3}{9} \times \frac{2}{8} & = & {}_3C_2 \times \frac{3 \times 2 \times 7}{10 \times 9 \times 8} \\ & & & & & = & \frac{7}{40} = 0.175 \end{array}$$

↓ もし、1000本中300本当たりのくじだったら?

$$(1) {}_3C_2 \left(\frac{300}{1000}\right)^2 \times \frac{700}{1000} = \frac{189}{1000} = 0.189$$

$$(2) {}_3C_2 \times \frac{300 \times 299 \times 700}{1000 \times 999 \times 998} \doteq 0.18893643$$

母集団の大きさが標本の大きさよりも十分に大きい場合、復元抽出と非復元抽出の違いは小さくなる。

したがって、非復元抽出による標本も近似的に

復元抽出による標本とみなすことができ、

$X_1, X_2, \dots, X_n$ はそれぞれが母集団分布に従う

互いに独立な確率変数と考えよう。

(2) 標本平均とその分布

(i) 標本平均・標本比率と正規分布

(ア) 標本平均の分布

例1 1, 2の数字が書いてあるカードが50枚ずつある。

この合計100枚のカードからなる母集団から、

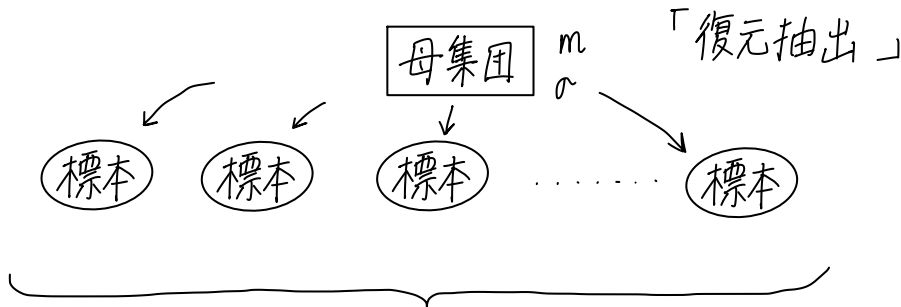
復元抽出によって大きさ3の無作為標本を抽出し、

そのカードの数字を順に X_1, X_2, X_3 とする。

このとき、標本平均 $\bar{X} = \frac{X_1 + X_2 + X_3}{3}$ について。

母集団			
標本	標本	標本	標本
X_1, X_2, X_3	X_1, X_2, X_3	X_1, X_2, X_3	X_1, X_2, X_3
1 1 1	1 1 2	1 2 2	2 2 2
$\bar{X} = \frac{1+1+1}{3}$	$\bar{X} = \frac{1+1+2}{3}$	$\bar{X} = \frac{1+2+2}{3}$	$\bar{X} = \frac{2+2+2}{3}$
$= 1$	$= \frac{4}{3}$	$= \frac{5}{3}$	$= 2$
$P = \left(\frac{1}{2}\right)^3$	$P = {}_3C_2 \left(\frac{1}{2}\right)^2 \times \frac{1}{2}$	$P = {}_3C_1 \times \frac{1}{2} \times \left(\frac{1}{2}\right)^2$	$P = \left(\frac{1}{2}\right)^3$
$= \frac{1}{8}$	$= \frac{3}{8}$	$= \frac{3}{8}$	$= \frac{1}{8}$

$$\text{よって、} E(\bar{X}) = 1 \times \frac{1}{8} + \frac{4}{3} \times \frac{3}{8} + \frac{5}{3} \times \frac{3}{8} + 2 \times \frac{1}{8} = \frac{3}{2} //$$



標本平均の平均 $E(\bar{X})$ を求める.

$$\begin{aligned}
 E(\bar{X}) &= E\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right) \\
 &= \frac{1}{n} E(X_1 + X_2 + \dots + X_n) \\
 &= \frac{1}{n} \{E(X_1) + E(X_2) + \dots + E(X_n)\} \\
 &= \frac{1}{n} \times m \times n = m
 \end{aligned}$$

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

まとめ 母平均 m , 母標準偏差 σ の母集団から
 大きさ n の無作為標本を抽出するとき、その標本平均 $\bar{X}$
 の期待値 $E(\bar{X})$ と標準偏差 $\sigma(\bar{X})$ は

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

例2 母平均10, 母分散9の母集団から,
大きさ4の標本を無作為に復元抽出したとき,
その標本平均の平均と分散, 標準偏差を求めよ。

解) 標本平均を $\bar{X}$ とすると,

$$E(\bar{X}) = 10 //$$

$$V(\bar{X}) = \frac{9}{4} //$$

$$\sigma(\bar{X}) = \sqrt{\frac{9}{4}} = \frac{3}{2} //$$

母平均 m , 母標準偏差 σ の母集団から抽出された
大きさ n の無作為標本について、標本平均 $\bar{X}$ は
近似的に正規分布 $N(m, \frac{\sigma^2}{n})$ に従うとみなすことができる。

$$\downarrow Z = \frac{\bar{X} - m}{\frac{\sigma}{\sqrt{n}}}$$

Z は近似的に標準正規分布 $N(0, 1)$ に従う。

例3 母平均が48、母標準偏差が16である母集団から
大きさ400の標本を無作為抽出するとき、その標本平均
を $\bar{X}$ とする。このとき、確率 $P(\bar{X} \geq 49)$ を求めよ。

標本平均 $\bar{X}$ は近似的に正規分布 $N(m, \frac{\sigma^2}{n})$ に従うとみなすことができる。

解)

母平均を m 、母標準偏差を σ 、標本の大きさを n とすると、

$$m = 48, \sigma = 16, n = 400$$

よって標本平均 $\bar{X}$ は近似的に正規分布 $N(48, \frac{16^2}{400})$

すなわち $N(48, 0.8^2)$ に従う。

$$\text{よって } P(\bar{X} \geq 49) = P(Z \geq \frac{49-48}{0.8})$$

$$= P(Z \geq 1.25)$$

$$= 0.5 - P(0 \leq Z \leq 1.25)$$

$$= 0.1056 //$$

(2) 標本平均とその分布

(i) 標本平均・標本比率と正規分布

(1) 標本比率と正規分布



抽出された標本の中での
ある特性Aをもつものの割合

母比率

母集団の中での
ある特性Aをもつものの割合

特性Aの母比率が p である、十分に大きな母集団から

大きさ n の無作為標本を抽出し、それらに対して

$X_1, X_2, X_3, \dots, X_n$ を次のように定める。

$$\begin{cases} \text{特性Aをもつとき} & X_R = 1 \\ \text{特性Aをもたないとき} & X_R = 0 \end{cases} \quad (R = 1, 2, \dots, n)$$

このとき $T = X_1 + X_2 + \dots + X_n$ を考えると、

T は大きさ n の標本の中での特性Aをもつものの個数

を表す確率変数であり、二項分布 $B(n, p)$ に従う。

また、標本平均 $\bar{X} = \frac{T}{n}$ は特性Aの標本比率 R を表す。

よって、 $q = 1 - p$ とすると、 n が大きいとき、

T は近似的に正規分布 $N(np, npq)$ に従う。

このとき $R = \frac{T}{n}$ は近似的に正規分布 $N\left(\frac{np}{n}, \frac{npq}{n^2}\right)$

に従う。

まとめ

特性 A の母比率 p の母集団から抽出された、大きさ n の無作為標本について、標本比率 R は、 n が大きいとき、近似的に正規分布 $N(p, \frac{pq}{n})$ に従うとみなすことができる。

例4

不良品が全体の10%に含まれる大量の製品の山から大きさ100の無作為標本を抽出するとき、不良品の標本比率 R について、次の問いに答えよ。

(1) R は近似的にどのような正規分布に従うとみなすことができるか。

(2) $0.07 \leq R \leq 0.13$ となる確率を求めよ。

解) (1) 母比率0.1の母集団から、大きさ100の

無作為標本を抽出するとき、標本比率 R は

近似的に正規分布 $N(0.1, \frac{0.1 \times 0.9}{100})$

すなわち $N(0.1, 0.03^2)$ に従う。

(2) $Z = \frac{R - 0.1}{0.03}$ とおくと、 Z は近似的に
標準正規分布 $N(0, 1)$ に従う。

$$\text{よって } P(0.07 \leq R \leq 0.13)$$

$$= P\left(\frac{0.07 - 0.1}{0.03} \leq Z \leq \frac{0.13 - 0.1}{0.03}\right)$$

$$= P(-1 \leq Z \leq 1)$$

$$= 2P(0 \leq Z \leq 1)$$

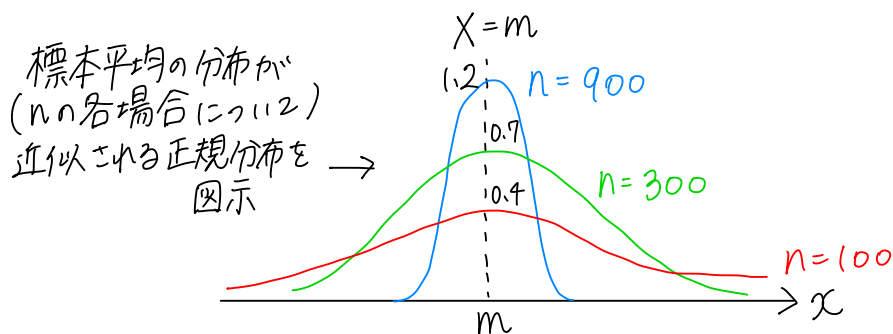
$$= 2 \times 0.3413 = 0.6826 //$$

(ii) 大数の法則

(復習)

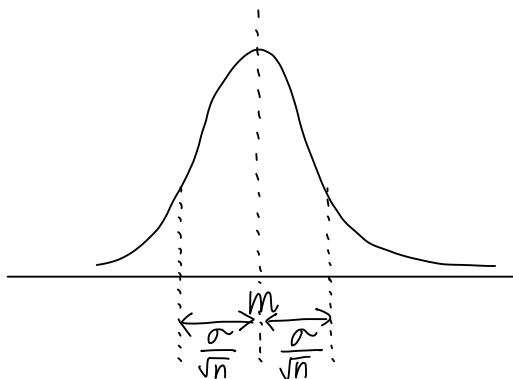
母平均 m , 母標準偏差 σ の母集団から抽出された
大きさ n の無作為標本について、標本平均 $\bar{X}$ は
近似的に正規分布 $N\left(m, \frac{\sigma^2}{n}\right)$ に従うとみなすことができる。

たとえば、 $\sigma = 10$ とし、 $n = 100, 300, 900$ としていくと、



大数の法則

母平均 m の母集団から大きさ n の無作為標本を抽出するとき、
 n が大きくなるに従って、標本平均 $\bar{X}$ は母平均 m に近づく



例1 母平均 0, 母標準偏差 1 の母集団から抽出した
大きさ n の標本の標本平均 $\bar{X}$ が -0.1 以上 0.1 以下
である確率を、 $n = 100, 400, 900$ の各場合について求めよ。

解) $Z = \frac{\bar{X} - 0}{\frac{1}{\sqrt{n}}} = \sqrt{n} \bar{X}$ とおくと、

Z は近似的に標準正規分布 $N(0, 1)$ に従う。

$$P(-0.1 \leq \bar{X} \leq 0.1) = P(-0.1 \times \sqrt{n} \leq Z \leq 0.1 \times \sqrt{n})$$

○ $n = 100$ のとき

$$\begin{aligned} P(-0.1 \leq \bar{X} \leq 0.1) &= P(-0.1 \times 10 \leq Z \leq 0.1 \times 10) \\ &= P(-1 \leq Z \leq 1) \\ &= 2P(0 \leq Z \leq 1) = 0.6826 // \end{aligned}$$

○ $n = 400$ のとき

$$\begin{aligned} P(-0.1 \leq \bar{X} \leq 0.1) &= P(-0.1 \times 20 \leq Z \leq 0.1 \times 20) \\ &= P(-2 \leq Z \leq 2) \\ &= 2P(0 \leq Z \leq 2) = 0.9544 // \end{aligned}$$

○ $n = 900$ のとき

$$\begin{aligned} P(-0.1 \leq \bar{X} \leq 0.1) &= P(-0.1 \times 30 \leq Z \leq 0.1 \times 30) \\ &= P(-3 \leq Z \leq 3) \\ &= 2P(0 \leq Z \leq 3) = 0.9973 // \end{aligned}$$

例2 1枚の硬貨を n 回投げるとき、表の出る相対度数を R とする。次の各場合について、確率 $P(|R - \frac{1}{2}| \leq 0.05)$ の値を求めよ。

(1) $n=100$ (2) $n=400$ (3) $n=900$

解)

相対度数(標本比率) R は、近似的に正規分布
 $N(\frac{1}{2}, \frac{\frac{1}{2} \cdot \frac{1}{2}}{n})$ すなわち $N(\frac{1}{2}, \frac{1}{4n})$ に従うので、

$Z = \frac{R - \frac{1}{2}}{\sqrt{\frac{1}{4n}}}$ は近似的に標準正規分布 $N(0, 1)$ に従う。

$$\begin{aligned} \therefore P(|R - \frac{1}{2}| \leq 0.05) &= P(|\sqrt{\frac{1}{4n}} Z| \leq 0.05) \\ &= P(|Z| \leq 0.05 \times \sqrt{4n}) \end{aligned}$$

(1) $n=100$ のとき

$$\begin{aligned} P(|R - \frac{1}{2}| \leq 0.05) &= P(|Z| \leq 0.05 \times 20) \\ &= P(|Z| \leq 1) \\ &= 2P(0 \leq Z \leq 1) = 0.6826 \end{aligned}$$

(2) $n = 400$ のとき

$$\begin{aligned}P\left(\left|R - \frac{1}{2}\right| \leq 0.05\right) &= P(|Z| \leq 0.05 \times 40) \\&= P(|Z| \leq 2) \\&= 2P(0 \leq Z \leq 2) = 0.9544\end{aligned}$$

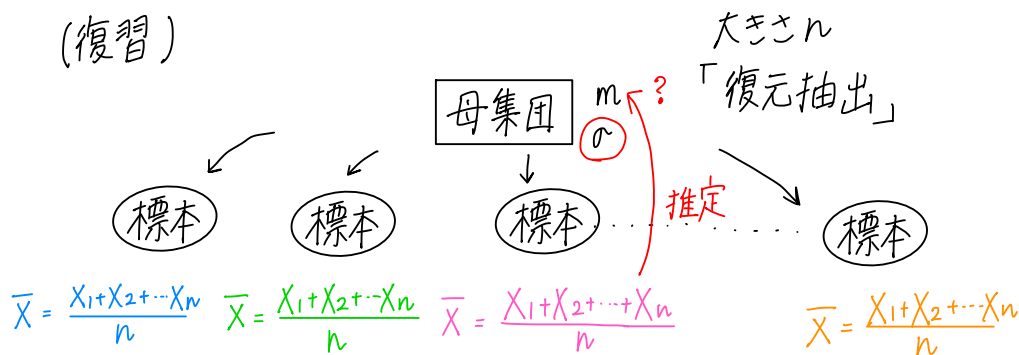
(3) $n = 900$ のとき

$$\begin{aligned}P\left(\left|R - \frac{1}{2}\right| \leq 0.05\right) &= P(|Z| \leq 0.05 \times 60) \\&= P(|Z| \leq 3) \\&= 2P(0 \leq Z \leq 3) = 0.9973\end{aligned}$$

(3) 推定

(i) 母平均の推定

(復習)



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

例) 日本人男性100人の身長を言ったらその平均が173cmであった。

全ての日本人男性
の平均身長も173cm
とする

「点推定」

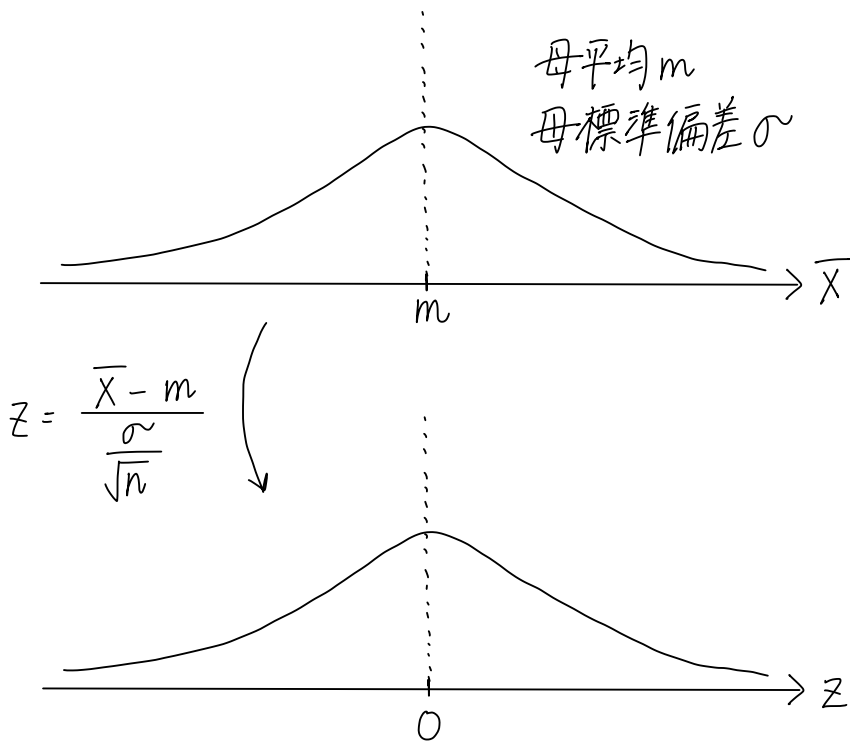
全ての日本人男性
の平均身長は171~175cm
に入るだろう

「区間推定」

信頼度95%の信頼区間

標本調査の結果から信頼区間を求めたとき、

100回のうち95回程度その区間に母平均が入るであろう。



$$P(|z| \leq 1.96) = 0.95$$

$$\rightarrow \left| \frac{\bar{X} - m}{\frac{\sigma}{\sqrt{n}}} \right| \leq 1.96$$

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

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

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

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

信頼度95%の信頼区間

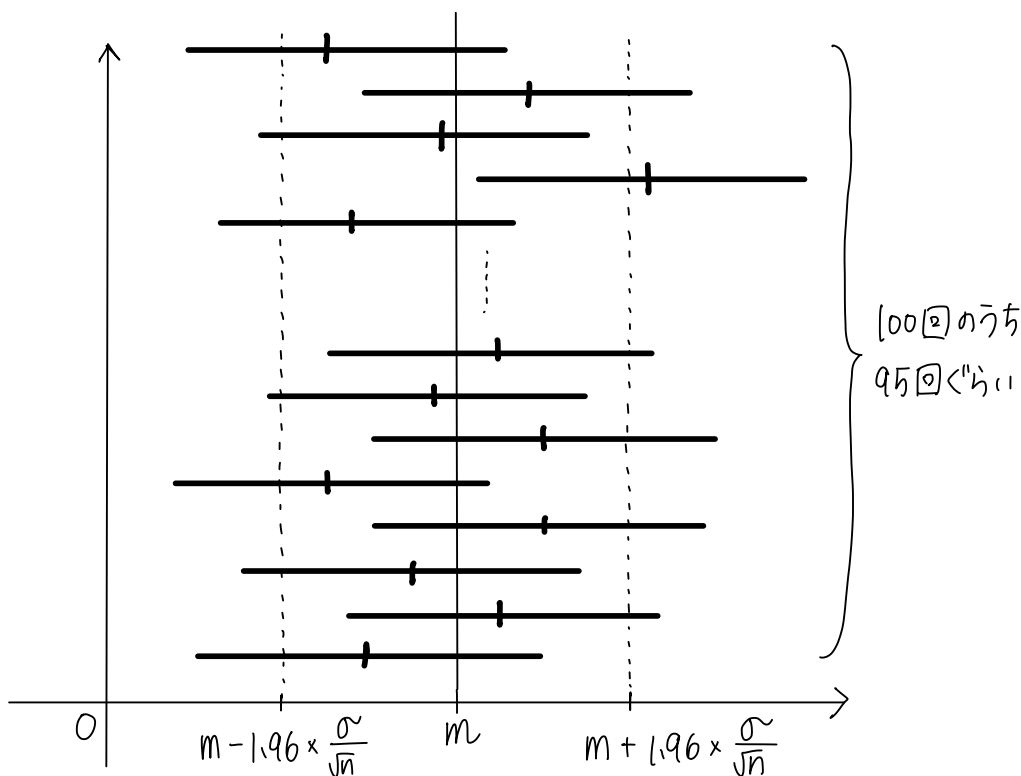
母平均の推定

標本の大きさが大きいとき、母平均 m に対する
信頼度 95% の信頼区間は

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

$$\text{または } \left[\bar{X} - 1.96 \times \frac{\sigma}{\sqrt{n}}, \bar{X} + 1.96 \times \frac{\sigma}{\sqrt{n}} \right]$$

標本調査の結果から信頼区間を求めたとき、
100回のうち95回程度その区間に母平均が入るのである。



例1

母標準偏差 20 の母集団から、大きさ 100 の標本を抽出したときの標本平均が 80 であった。このとき、母平均 m に対する信頼度 95% の信頼区間を求めよ。

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

解)

母標準偏差が 20、標本の大きさが 100、
標本平均が 80 なの2、母平均 m に対する
信頼度 95% の信頼区間は

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

$$80 - 1.96 \times \frac{20}{\sqrt{100}} \leq m \leq 80 + 1.96 \times \frac{20}{\sqrt{100}}$$

$$76.08 \leq m \leq 83.92$$

したがって、76.1 以上 83.9 以下である。

標本の大きさが大きいとき、母平均 m に対する信頼度

$$95\% \text{ の信頼区間は、 } \bar{X} - 1.96 \times \frac{\sigma}{\sqrt{n}} \leq m \leq \bar{X} + 1.96 \times \frac{\sigma}{\sqrt{n}}$$

$$99\% \text{ の信頼区間は、 } \bar{X} - 2.58 \times \frac{\sigma}{\sqrt{n}} \leq m \leq \bar{X} + 2.58 \times \frac{\sigma}{\sqrt{n}}$$

母標準偏差 σ がわからないときは

σ のかわりに標本標準偏差 S を用いよ!

例2 ある工場で生産された製品の中から、100個を無作為に抽出して重さを量ったところ、平均 50.4g、標準偏差 1.53g であった。この製品の平均の重さ m g に対して、信頼度 95% の信頼区間を求めよ。

解)

標本の大きさが 100、標本平均が 50.4、
標本の標準偏差が 1.53 なの2、母平均 m に対する
信頼度 95% の信頼区間は

$$50.4 - 1.96 \times \frac{1.53}{\sqrt{100}} \leq m \leq 50.4 + 1.96 \times \frac{1.53}{\sqrt{100}}$$

$$\text{よって、} 50.1 \leq m \leq 50.7$$

したがって、この製品の平均の重さ m g に対して、
信頼度 95% の信頼区間は $[50.1, 50.7]$

(ii) 母比率の推定

(復習)

特性 A の母比率 p の母集団から抽出された、大きさ n の無作為標本について、標本比率 R は、 n が大きいとき、

近似的に正規分布 $N(p, \frac{p(1-p)}{n})$ に従うとみなすことができる。

母比率の推定

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

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

(例3)

ある都市での世論調査において、無作為に 400 人の有権者を選び、ある政策に対する賛否を調べたところ、320 人が賛成であった。この都市の有権者全体のうち、この政策の賛成者の母比率 p に対する信頼度 95% の信頼区間を求めたい。

(1) 標本比率 R を求めよ。

(2) p に対する信頼度 95% の信頼区間を求めよ。

(小数第 2 位まで)

解)

(1) 標本比率 R は $\frac{320}{400} = \frac{4}{5} = 0.8 //$

(2) 標本の大きさは 400 であるから、

母比率 p に対する信頼度 95% の信頼区間は

$$\left[0.8 - 1.96 \times \sqrt{\frac{0.8 \times 0.2}{400}}, 0.8 + 1.96 \times \sqrt{\frac{0.8 \times 0.2}{400}} \right]$$

つまり、 $[0.76, 0.84] //$

(4) 仮説検定

例1 あるさいころを720回投げたところ、1の目が100回出た。
このさいころの1の目が出る確率は $\frac{1}{6}$ ではないと判断してよいか。

主張したい 「1の目が出る確率は $\frac{1}{6}$ ではない」

「1の目が出る確率は $\frac{1}{6}$ である」

帰無仮説が正しいかを言明していく。

解)

「1の目が出る確率は $\frac{1}{6}$ である」

が正しいとすると、720回のうち1の目が出る回数 X は

二項分布 $B(720, \frac{1}{6})$ に従う。

$$E(X) = 720 \times \frac{1}{6} = 120$$

$$\sigma(X) = \sqrt{720 \times \frac{1}{6} \times \frac{5}{6}} = 10$$

よって、 $Z = \frac{X - 120}{10}$ は標準正規分布 $N(0, 1)$ に従う

とみなしてよい。

正規分布表より $P(-1.96 \leq Z \leq 1.96) = 0.95$ である。

$X = 100$ のとき $Z = -2$

この値は棄却域に含まれ、帰無仮説は棄却できる。

このさいころの1の目が出る確率は $\frac{1}{6}$ ではないと判断できる。

仮説検定の流れ

有意水準5%の場合

① 対立仮説 (正しいかどうか判断したい仮説を立てる。

② 帰無仮説を立てる。

③ 帰無仮説が正しいと仮定したときの

実際のデータのことが起こる確率 P を求める。

$P < 0.05$

帰無仮説は棄却される。

対立仮説が正しいと
判断できる。

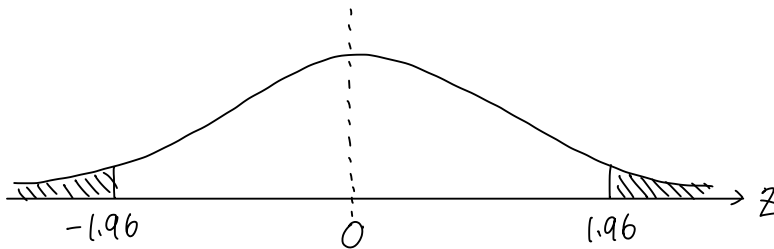
$P \geq 0.05$

帰無仮説は棄却されない

対立仮説が正しいかどうか
判断できない

両側検定

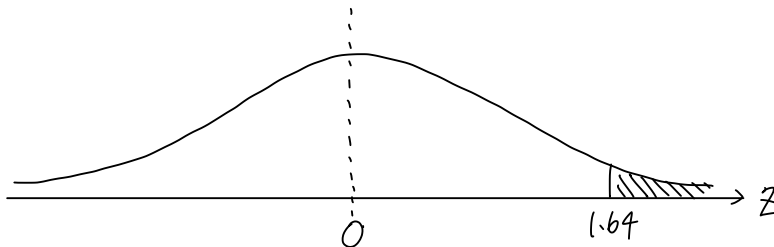
「ある値と異なる」



片側検定

「ある値より大きい」

「ある値より小さい」



例2

ある企業が自社製品の知名度を上げるため広告を出した。
 広告を出す前のアンケート調査で、製品を知っている人数は
 2500人中50人であったが、広告を出した後のアンケート調査
 では2500人中63人になった。この広告は効果があったと
 考えてよいか。有意水準5%で片側検定せよ。



解)

広告を出す前に製品を知っている人の割合は $\frac{50}{2500} = 0.02$

対立仮説を

「広告の後に製品を知っている人の割合が0.02より増えた」

帰無仮説を

「広告の後に製品を知っている人の割合が0.02である。」とする。

帰無仮説が正しいと仮定すると、2500人のうち、製品を知っている
 人数 X は二項分布 $B(2500, 0.02)$ に従う。

$$E(X) = 2500 \times 0.02 = 50$$

$$\sigma(X) = \sqrt{2500 \times 0.02 \times 0.98} = 7$$

よって、 $Z = \frac{X-50}{7}$ は標準正規分布 $N(0, 1)$ に従うものとみなせる。

$$X = 63 \text{ のとき } Z = \frac{13}{7} \doteq 1.857$$

正規分布表より $P(Z \geq 1.64) = 0.05$ である。

この値は棄却域に含まれ、帰無仮説は棄却できる。

したがって、この広告は効果があったと判断できる。