確率、確率変数、確率分布

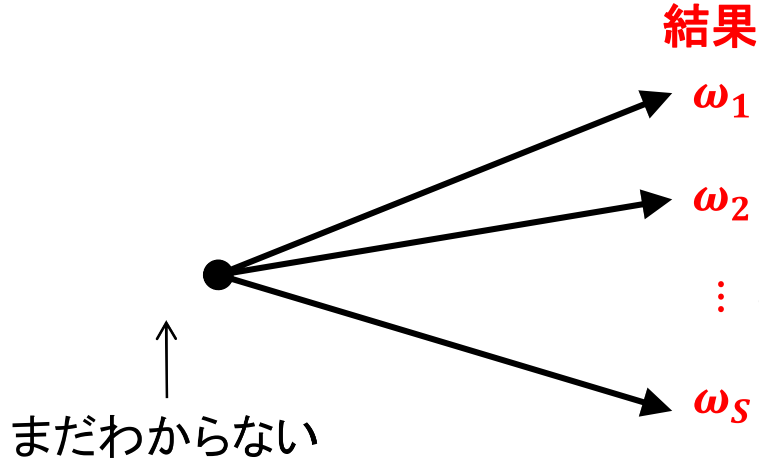
データ・マイニングI

確率

結果

- <u>**結果とは、不確実性が伴う現象や状況**</u>

$$\{\omega_1\},\quad \{\omega_2\},\quad \cdots,\quad \{\omega_S\}$$

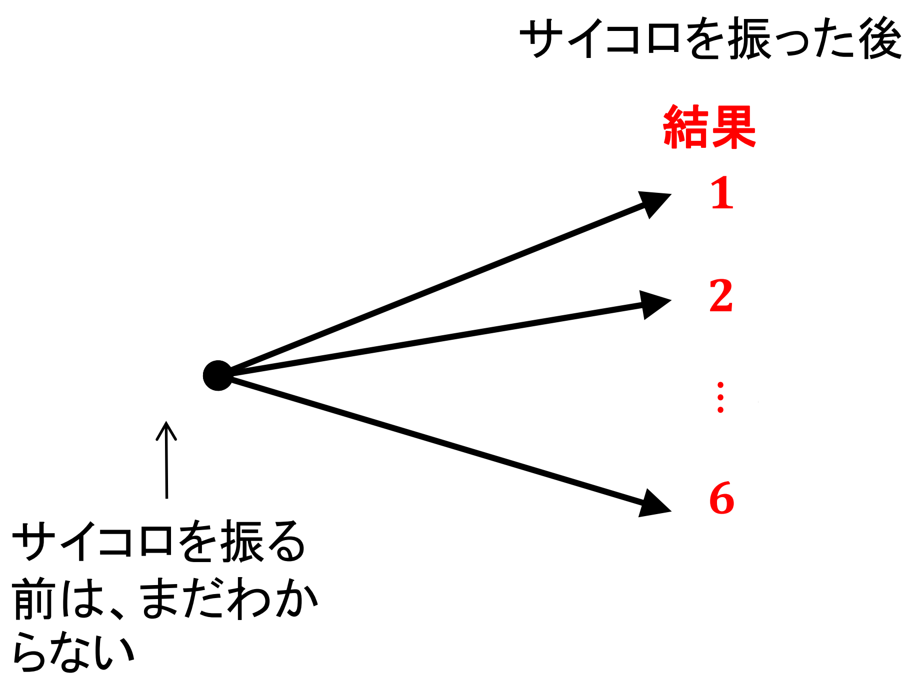


15:50:47

結果 : サイコロ

- **<u>結果とは、不確実性が伴う現象や状況</u>**

$$\{\mathbf{1}\},\quad \{\mathbf{2}\},\quad \{\mathbf{3}\},\quad \{\mathbf{4}\},\quad \{\mathbf{5}\},\quad \{\mathbf{6}\}$$



事象

- **事象とは、不確実性が伴う結果の集まり**

$\boldsymbol{\phi}$ = 空集合(何もない集合),

$\{\omega_1\}$, $\{\omega_2\}$, $\cdots$, $\{\omega_S\}$,

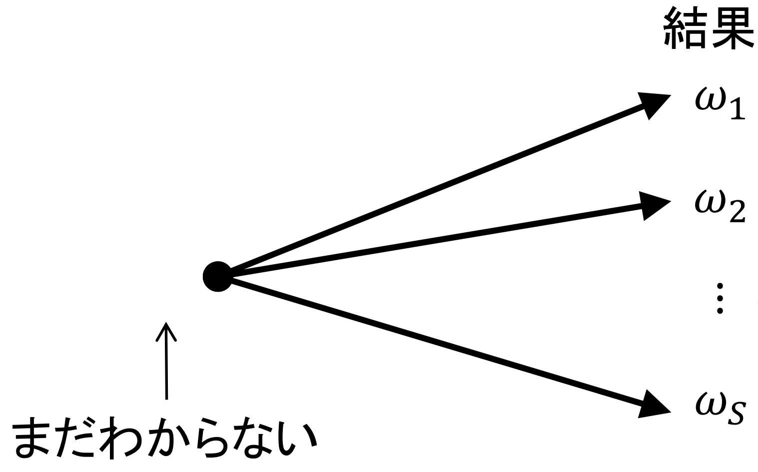
$\{\omega_1, \omega_2\}$, $\{\omega_1, \omega_3\}$, $\cdots$, $\{\omega_1, \omega_S\}$, $\cdots$,

$\{\omega_{S-1}, \omega_S\}$

$\{\omega_1, \omega_2, \omega_3\}$, $\cdots$, $\{\omega_{S-2}, \omega_{S-1}, \omega_S\}$

$\cdots$,

$\{\omega_1, \omega_2, \omega_3, \cdots, \omega_S\}$



15:50:48

事象の例：サイコロ

- **事象とは、不確実性が伴う結果の集まり**

ϕ = 空集合(何もない集合),

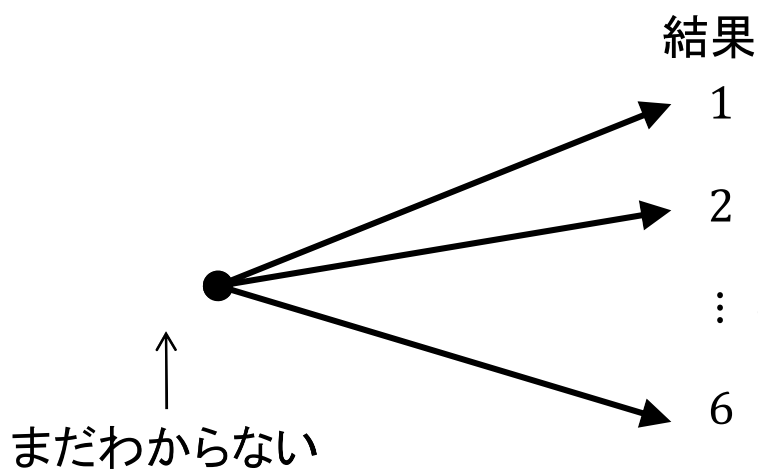
$\{1\},\quad \{2\},\quad \{3\},\quad \{4\},\quad \{5\},\quad \{6\},$

$\{1,2\},\quad \{1,3\},\quad \cdots,\quad \{1,6\},\quad \cdots,\quad \{5,6\}$

$\{1,2,3\},\quad \cdots,\quad \{4,5,6\}$

$\cdots,$

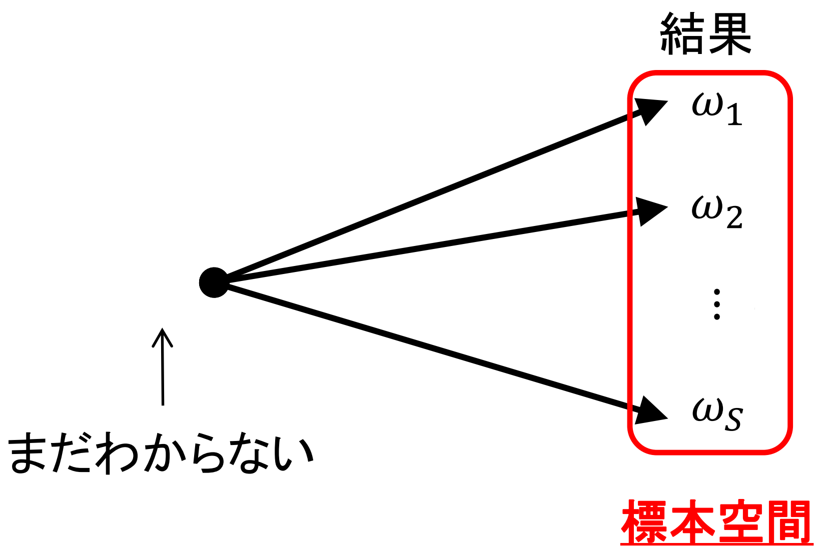
$\{1,2,3,4,5,6\}$



標本空間

- <u>**標本空間とは、結果全体の集合**</u>
  - <u>$\mathbf{\Omega}$で表す</u>

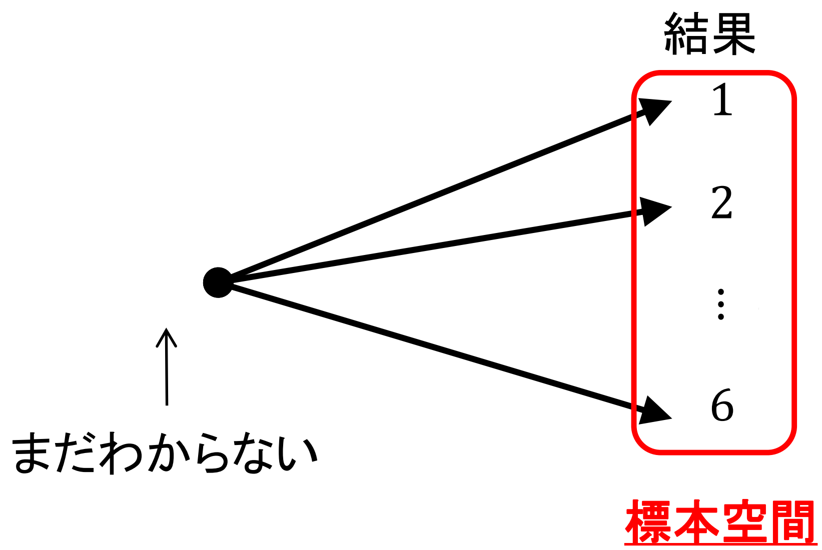
$$\boldsymbol{\Omega = \{\omega_1, \omega_2, \cdots, \omega_S\}}$$



標本空間の例：サイコロ

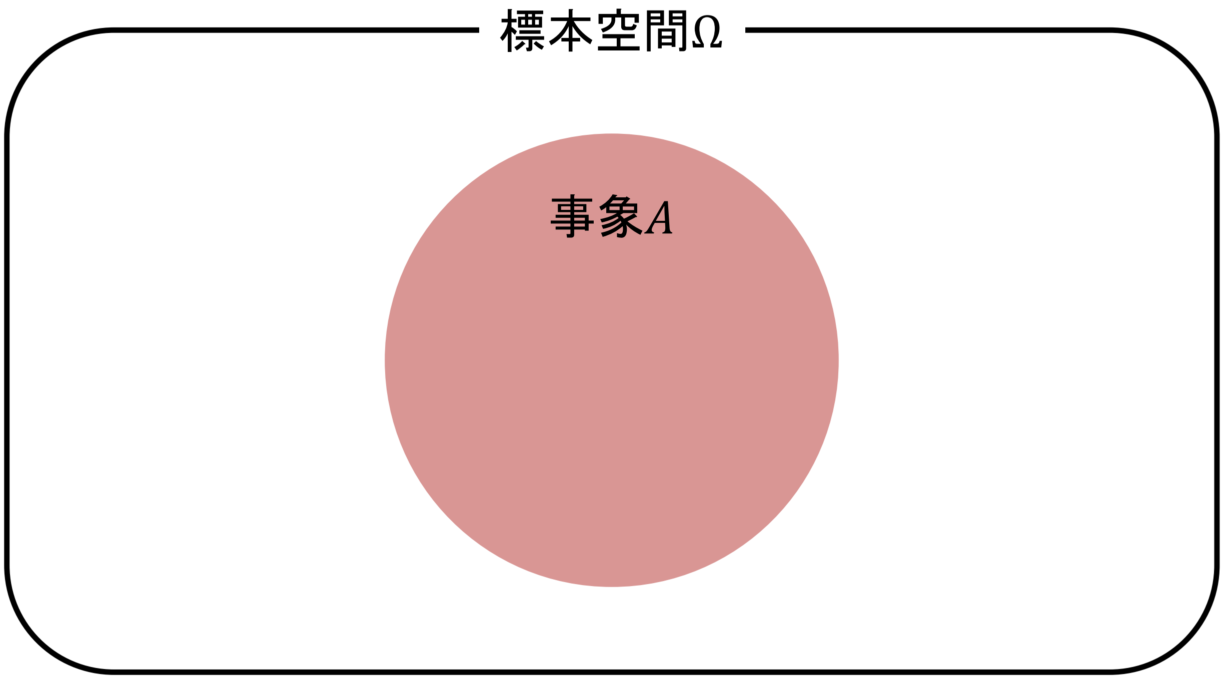
- **標本空間とは、結果全体の集合**

$$\boldsymbol{\Omega} = \{\mathbf{1}, \mathbf{2}, \mathbf{3}, \mathbf{4}, \mathbf{5}, \mathbf{6}\}$$



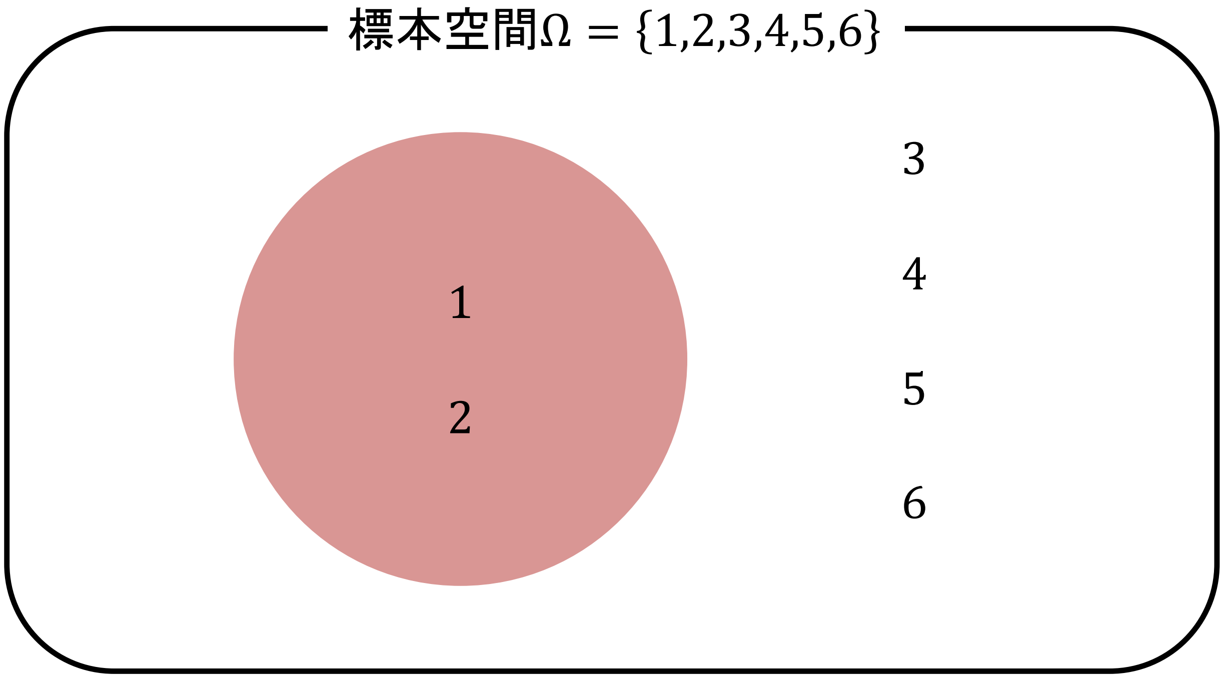
集合ーVenn図ー

- 標本空間Ωの中に事象Aがある



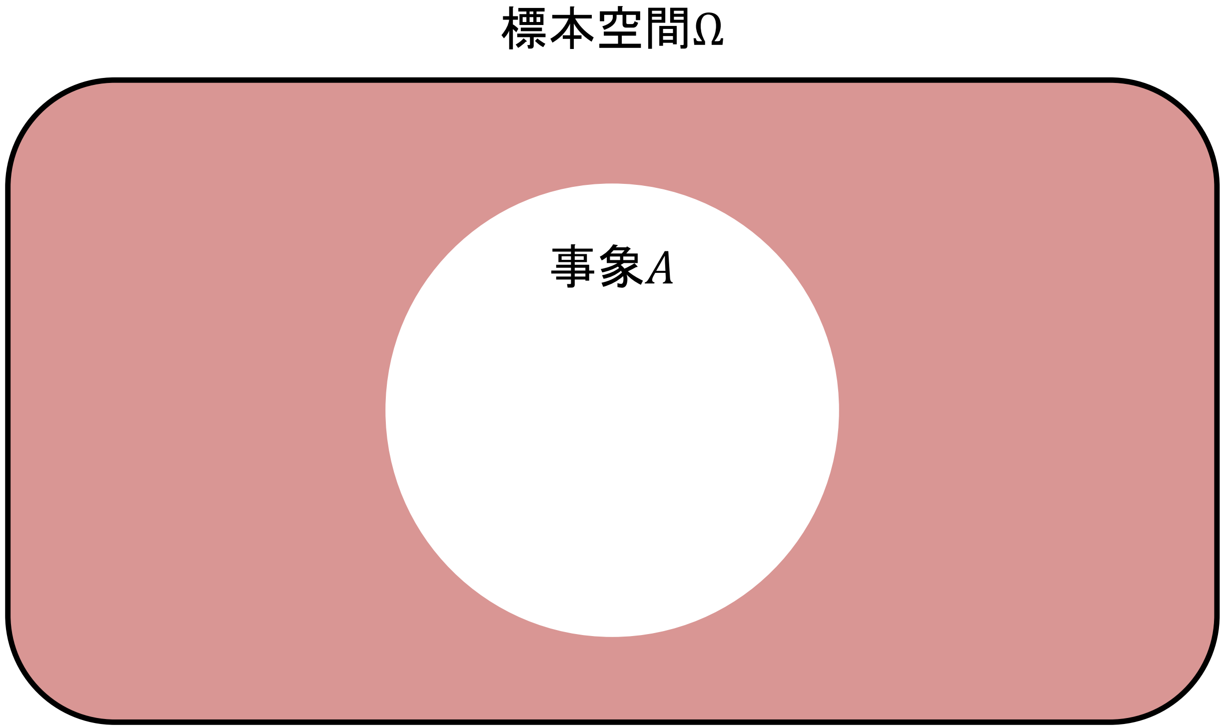
集合ーVenn図ー　例: サイコロの目

- 標本空間Ωの中に事象Aがある



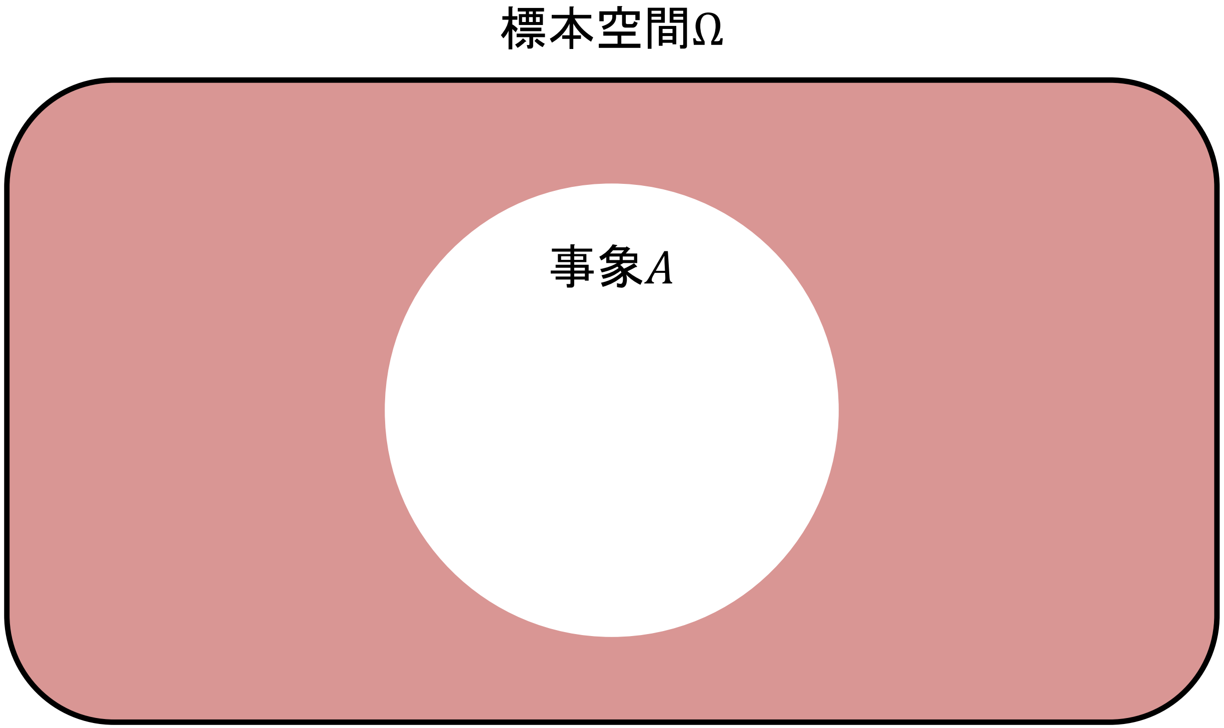
差集合と補集合

- **差集合$B-A$とは、集合Bに含まれ、かつ集合Aに含まれない点の集合**
  - 事象Bから事象Aを差し引いたもの
  - 特に、**標本空間Ωから事象Aを差し引いたものAの補集合とも言う**



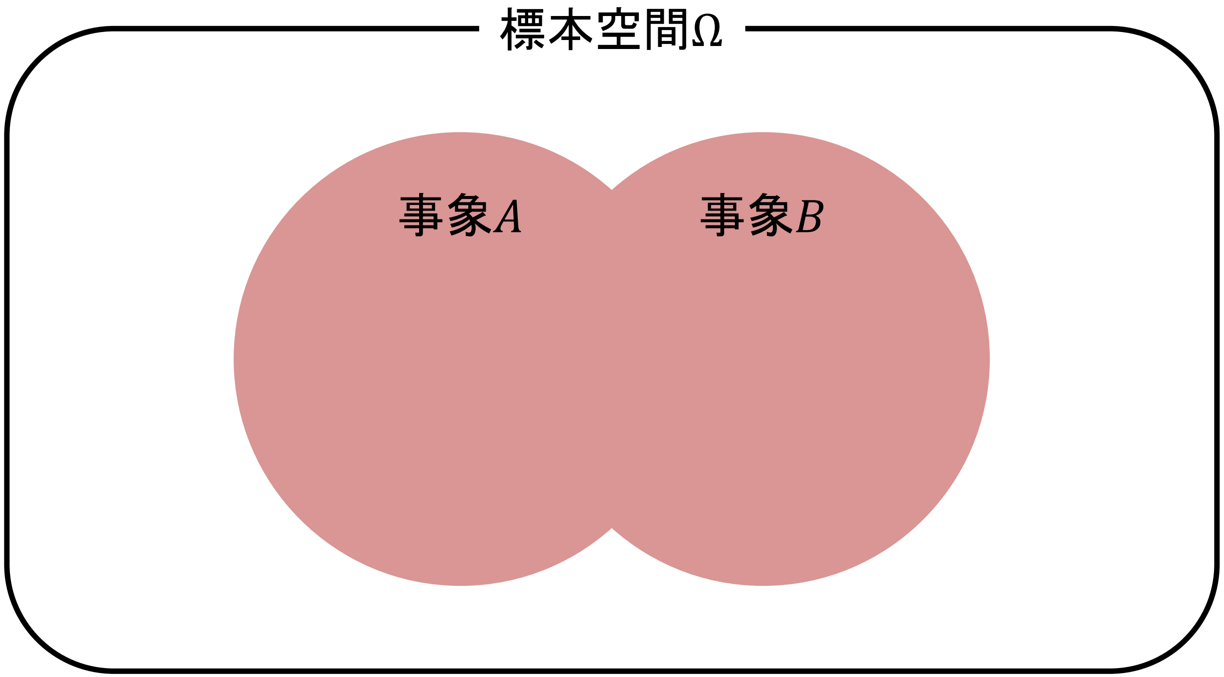
補集合

- **<u>差集合$\Omega - A$とは、標本空間Ωに含まれ、かつ集合Aに含まれない点の集合</u>**
  - 標本空間Ωから事象Aを差し引いたもの
  - 特に、**<u>標本空間Ωから事象Aを差し引いたものAの補集合とも言う</u>**



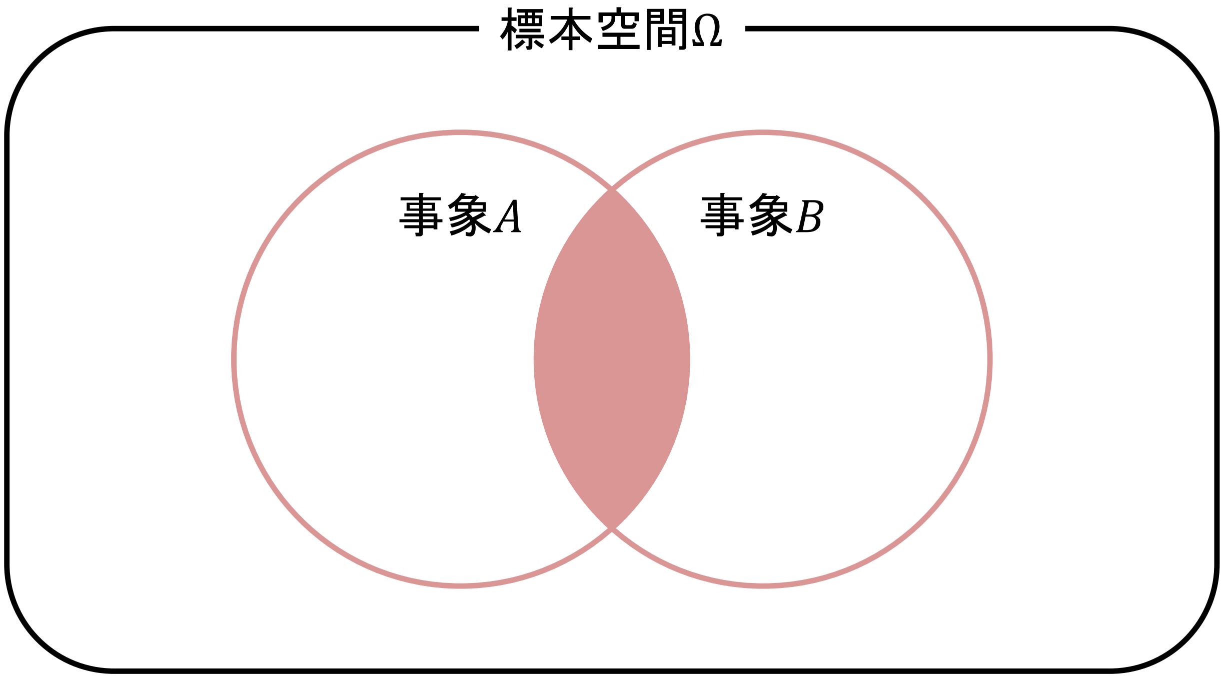
和集合

- **和集合$A \cup B$とはAもしくはBに含まれる点の集合**
  - 事象Aと事象Bを合わせたもの
  - 事象Aもしくは事象Bのいずれかが起こるときを示す



共通集合

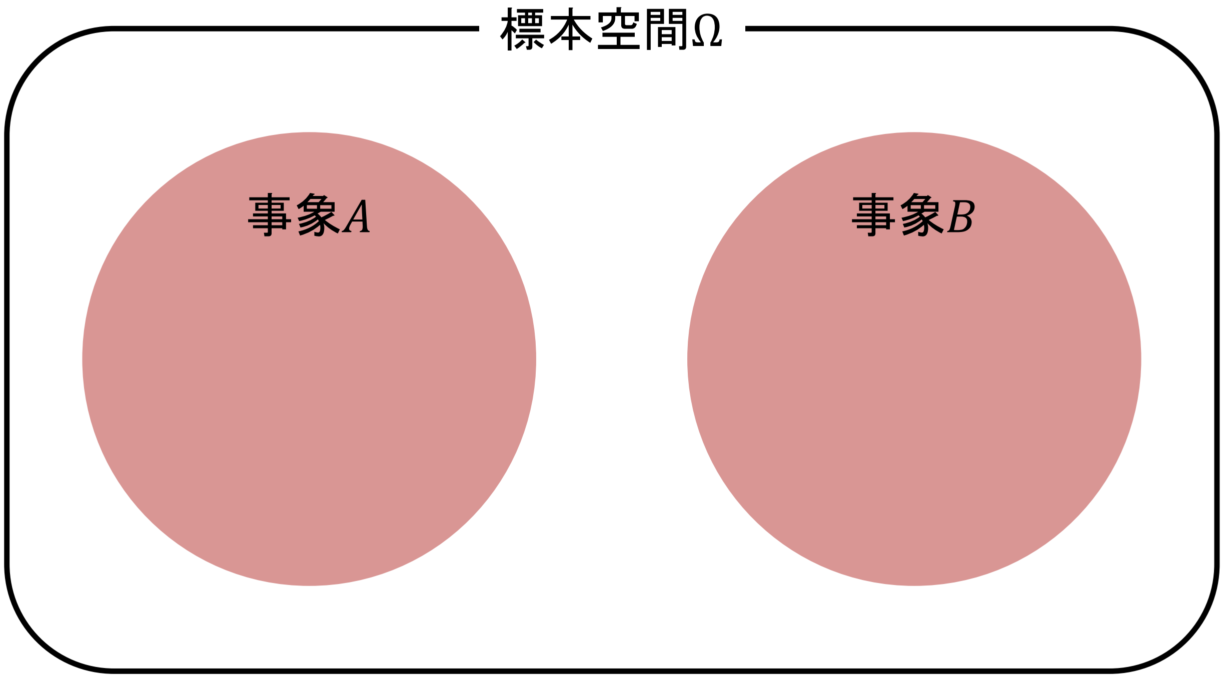
- **<u>共通集合$A \cap B$とはAかつBに含まれる点の集合</u>**
  - 事象Aと事象Bの共通部分
  - 事象Aと事象Bが同時に起こるときを表す



15:50:49

排反な集合

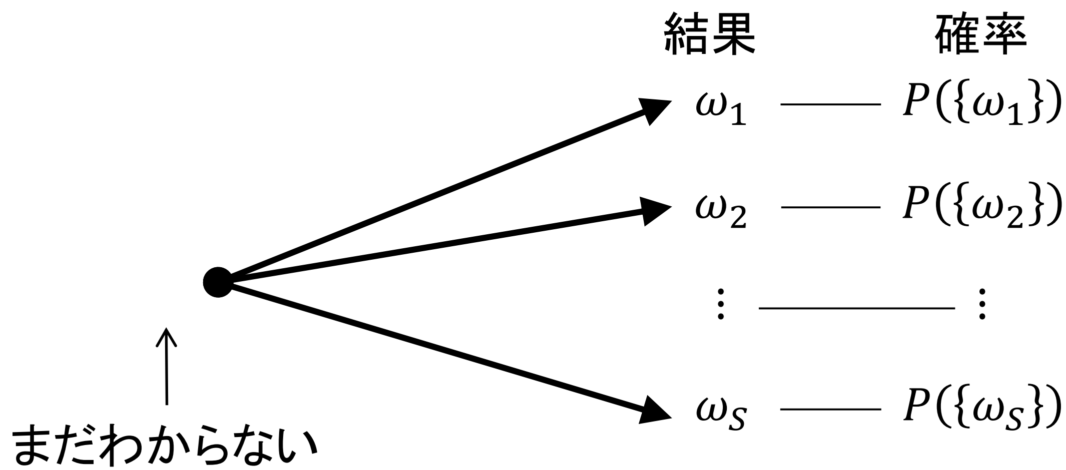
- <u>**共通集合$A \cap B$がないとき、AとBは互いに排反である**</u>
  - 共通集合$A \cap B = \phi$（空集合）
  - AとBが同時に起こることはない



事象の確率

- <u>**事象Aが生じる確率を$P(A)$と表す**</u>

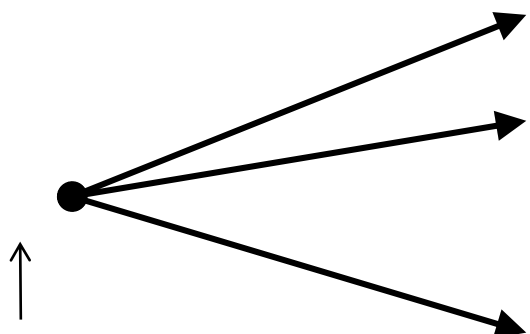
$$\boldsymbol{P(\phi)},$$
$$\boldsymbol{P(\{\omega_1\}),\quad P(\{\omega_1\}),\quad \cdots,\quad P(\{\omega_S\}),}$$
$$\boldsymbol{P(\{\omega_1,\omega_2\}),\quad P(\{\omega_1,\omega_3\}),\quad \cdots,\quad P(\{\omega_{S-1},\omega_S\})}$$
$$\boldsymbol{P(\{\omega_1,\omega_2,\omega_3\}),\quad \cdots,\quad P(\{\omega_{S-2},\omega_{S-1},\omega_S\})}$$
$$\cdots,$$
$$\boldsymbol{P(\{\omega_1,\omega_2,\omega_3,\cdots,\omega_S\})}$$



事象の確率の例: サイコロ

- <u>**事象Aが生じる確率を$P(A)$と表す**</u>

$$P(\phi),$$
$$P(\{1\}),\quad P(\{2\}),\quad P(\{3\}),\quad P(\{4\}),\quad P(\{5\}),$$
$$P(\{6\}),$$
$$P(\{1,2\}),\quad P(\{1,3\}),\quad \cdots,\quad P(\{5,6\})$$
$$P(\{1,2,3\}),\quad \cdots,\quad P(\{4,5,6\})$$
$$\cdots,$$
$$P(\{1,2,3,4,5,6\})$$



まだわからない

結果	確率
1	$P(\{1\}) = 1/6$
2	$P(\{2\}) = 1/6$
$\vdots$	$\vdots$
6	$P(\{6\}) = 1/6$

15:50:49

事象の確率の例：サイコロ

- **事象Aが生じる確率を$P(A)$と表す**

$$P(\phi) = 0$$

$$P(\{1\}) = \frac{1}{6}, \quad \cdots, \quad P(\{6\}) = \frac{1}{6},$$

$$P(\{1,2\}) = \frac{1}{3}, \quad \cdots, \quad P(\{5,6\}) = \frac{1}{3}$$

$$P(\{1,2,3\}) = \frac{1}{2}, \quad \cdots, \quad P(\{4,5,6\}) = \frac{1}{2}$$

$$\cdots,$$

$$P(\{1,2,3,4,5,6\}) = 1$$

確率の基本的な特徴

- すべての事象の確率は0から1の間を取る

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

- 空集合の確率は0、すべての結果の集合（標本空間）の確率は1

$$P(\phi) = 0, \qquad P(\Omega) = 1$$

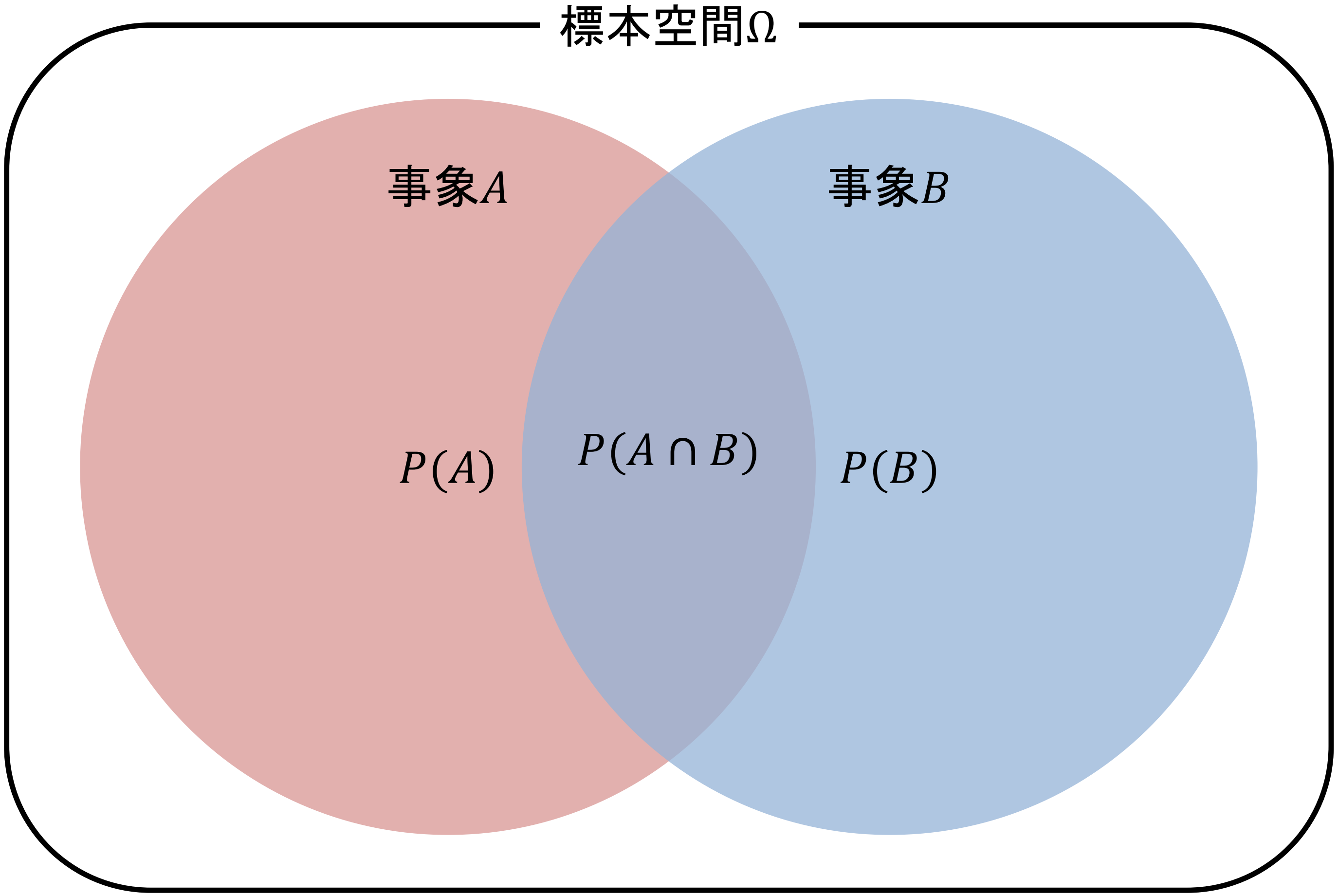
- すべての結果の確率を足し合わせると1

$$P(\{\omega_1\}) + P(\{\omega_2\}) \cdots + P(\{\omega_S\}) = 1$$

確率の計算：確率の足し算

- 事象Aと事象Bの和集合の確率は

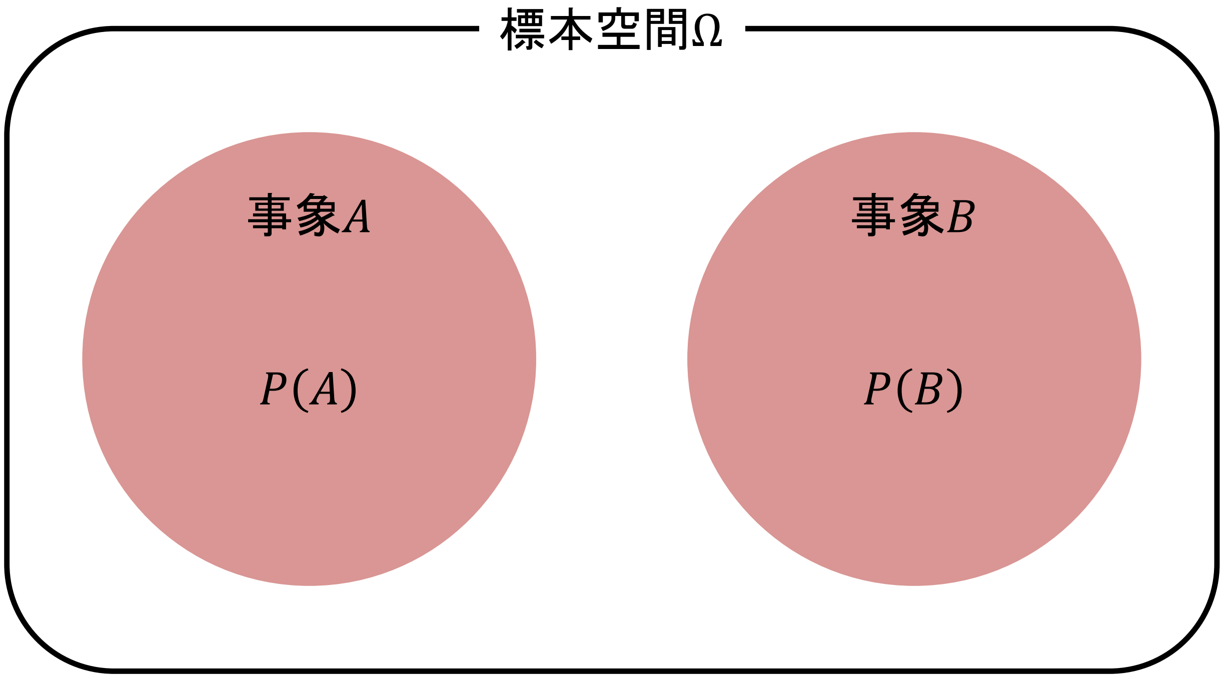
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$



排反な事象の確率の足し算

- 事象Aと事象Bが排反だった場合の和集合の確率は

$$P(A \cup B) = P(A) + P(B)$$

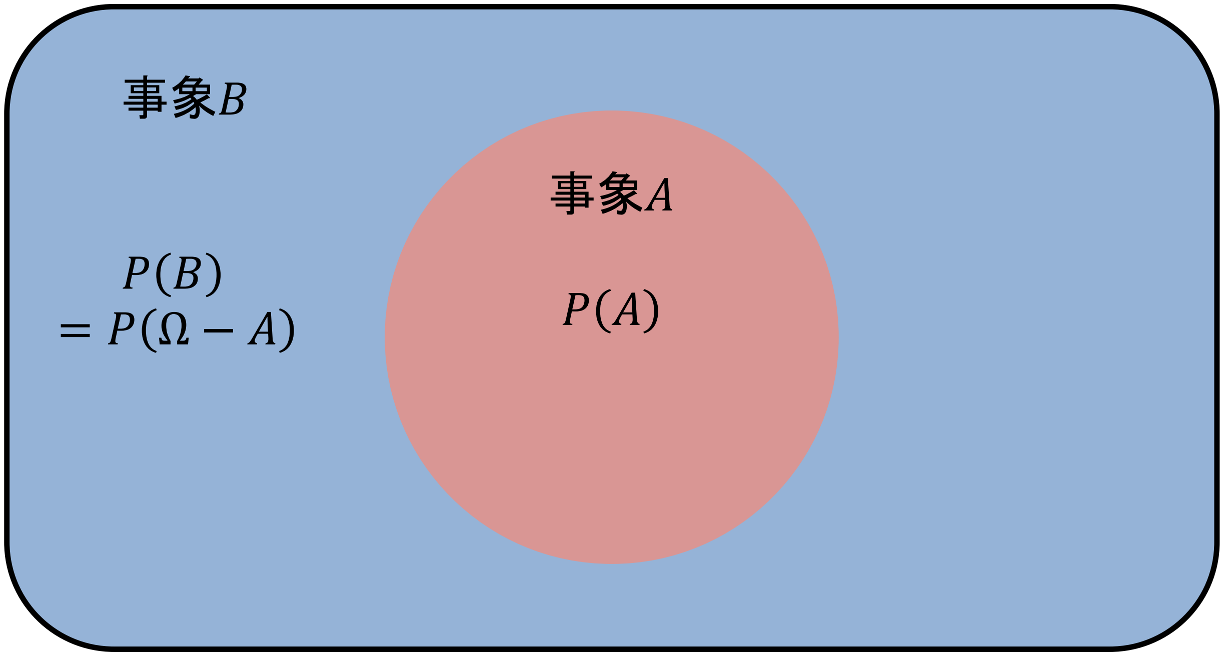


差集合に対する確率の計算

- 差集合の事象$B = \Omega - A$の確率は事象Aの確率から計算できる

$$P(B) = 1 - P(A)$$

標本空間Ω



15:50:50

相対的頻度

- 相対的頻度とは、

$$P(A) = \frac{\mathrm{fr(A)}}{\mathrm{N}} = \frac{\text{ある事象}A\text{の頻度}}{\text{生じた結果の全ての回数（度数）}}$$

  - サイコロを100回振った
  - 3の目が出た回数が17回だった
  - 相対頻度は

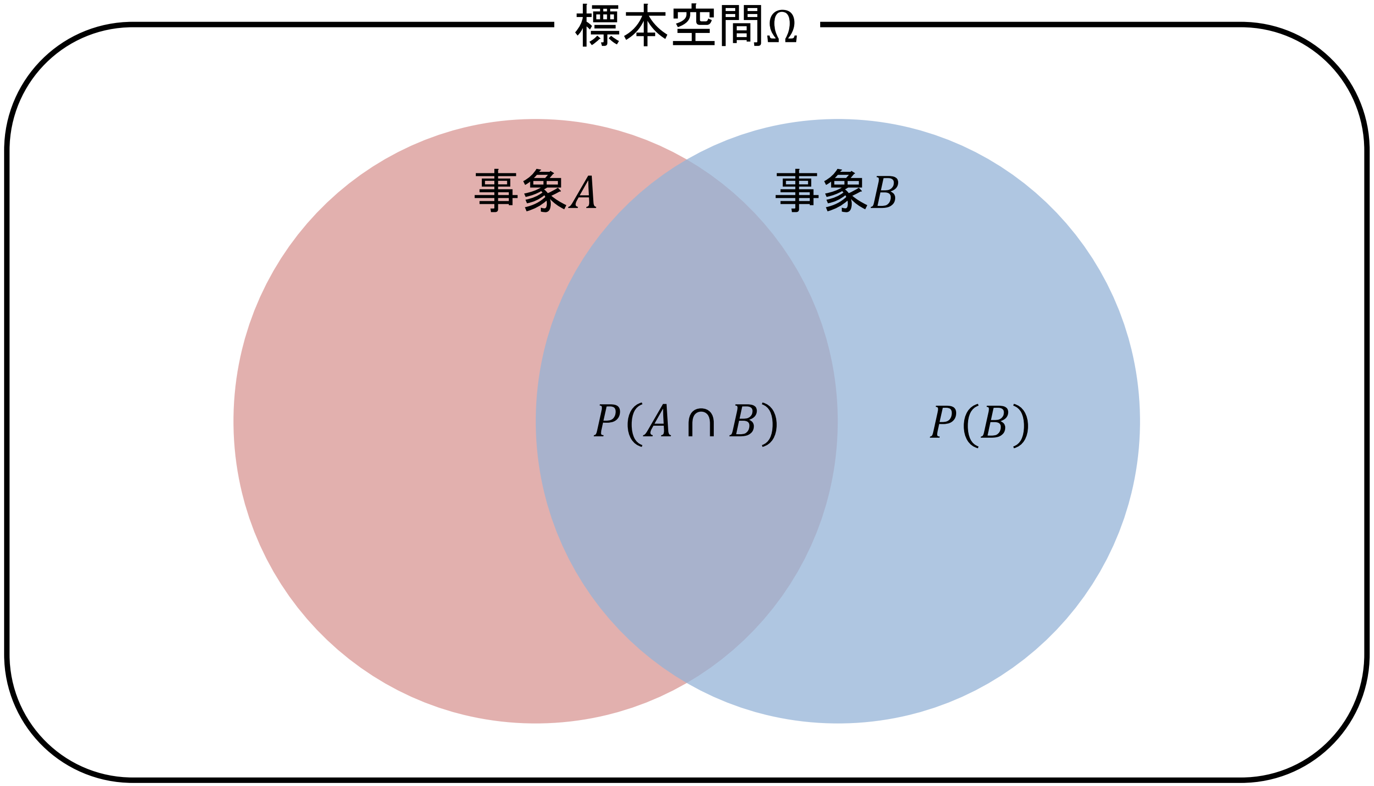
$$\frac{17}{100} = 0.17 = 17\%$$

  - 本当の確率ではなく、あくまでも実験した結果の頻度を割合として示したもの
    - 3の目が出る真の確率は$\frac{1}{6} \approx 0.1667$

確率の計算：条件付き確率

- $P(B) \neq 0$のとき、事象Bが与えられたときの事象Aの条件付き確率は

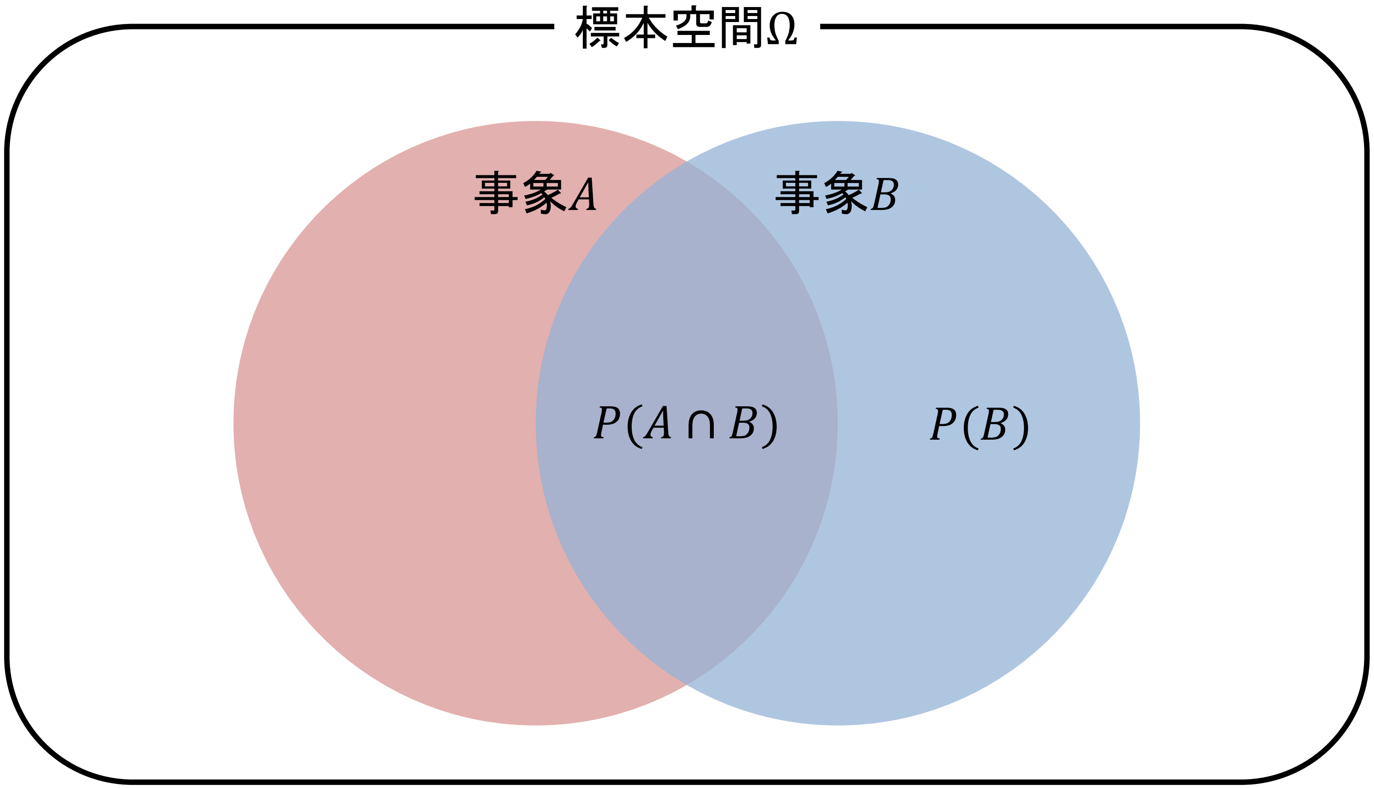
$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$



確率の計算：条件付き確率による共通集合の確率

- $P(B) \neq 0$のとき、事象Aかつ事象Bの確率は

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

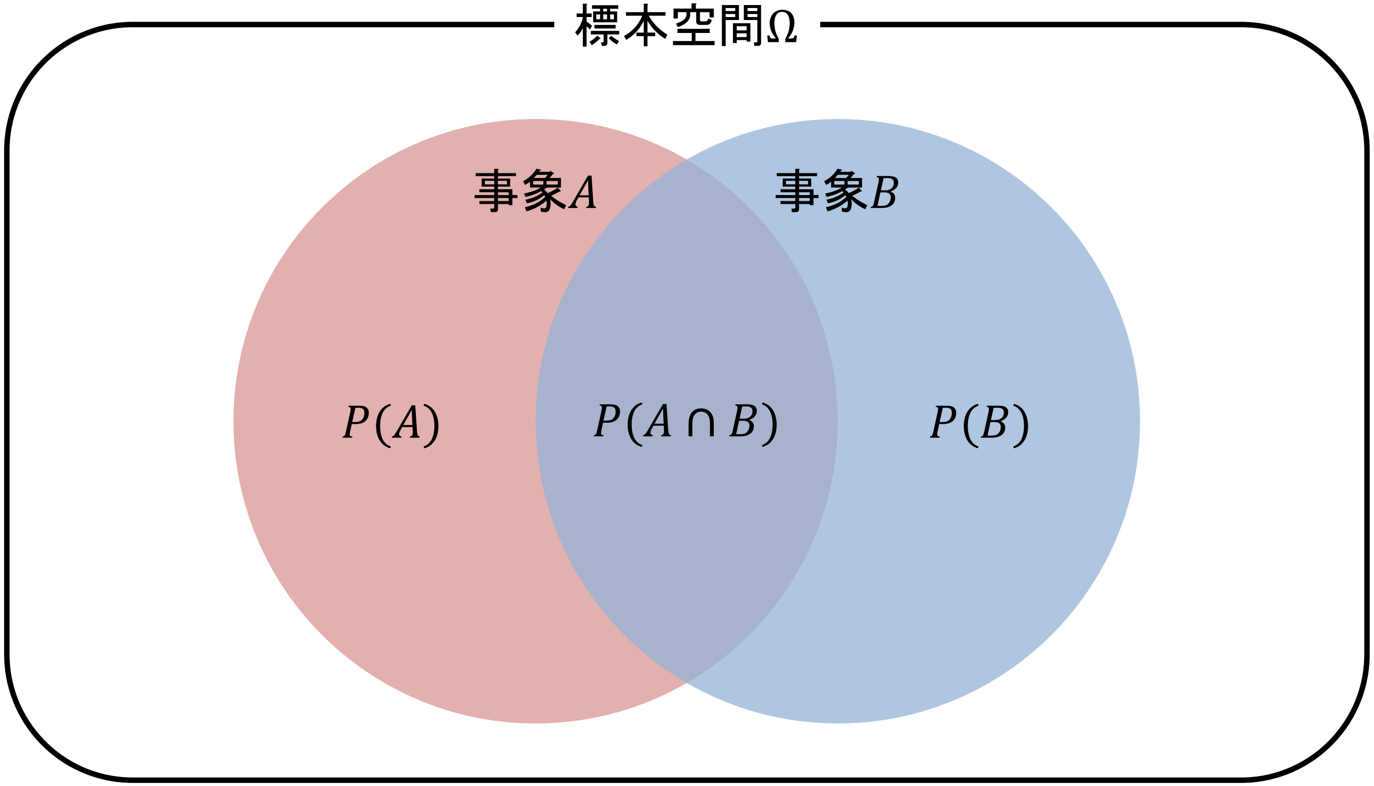


15:50:50

独立な事象

- **事象Aと事象Bは以下の条件を満たすとき独立であるという**

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



15:50:50

Bayesの定理

- **Bayesの公式は**

$$\boldsymbol{P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A')P(A')}}$$

ここで、事象$\boldsymbol{A}$と事象$\boldsymbol{A'}$は排反であり、$\boldsymbol{\Omega = A \cup A'}$

- $P(A|B)P(B) = P(A \cap B) = P(B|A)P(A)$より最初の等式が得られる

- $P(B) = P(B \cap A) + P(B \cap A') = P(B|A)P(A) + P(B|A')P(A')$
  - ここで、排反な事象の確率の足し算と条件付確率による共通事象の確率を用いる

15:50:50

拡張版Bayesの定理

- **Bayesの公式は**

$$\boldsymbol{P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{P(B|A)P(A)}{\sum_{\{i=1\}}^{I} P(B|A_i)P(A_i)}}$$

ここで、事象$\boldsymbol{A_1, A_2, \cdots, A_I}$は排反であり、$\boldsymbol{\Omega = A_1 \cup A_2 \cup \cdots \cup A_I}$

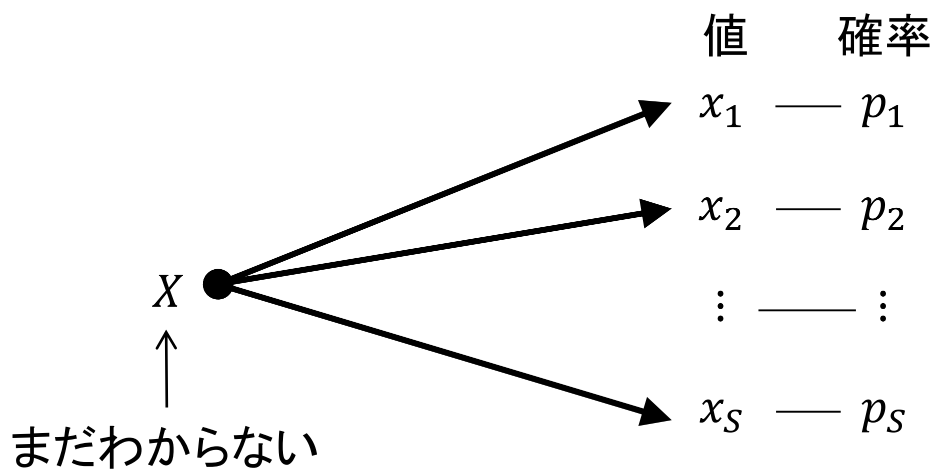
- $P(B) = P(B \cap A_1) + \cdots + P(B \cap A_I) = P(B|A_1)P(A_1) + \cdots + P(B|A_I)P(A_I)$
  - ここで、排反な事象の確率の足し算と条件付確率による共通事象の確率を用いる

YOUR TURN

確率変数

離散的な確率変数

1. **離散的な確率変数Xはs 個の数値$(x_1, x_2, \cdots, x_S)$の内いずれかを取る変数であり、**
2. **各値に対してその値が起こる確率$(p_1, p_2, \cdots, p_S)$が対応している**
   - 正確には$p_1 = P(\{X = x_1\}), p_2 = P(\{X = x_2\}), \ldots, p_S = P(\{X = x_S\})$
   - $P(\{X = x\})$は確率変数Xがあるxとなる確率を意味する

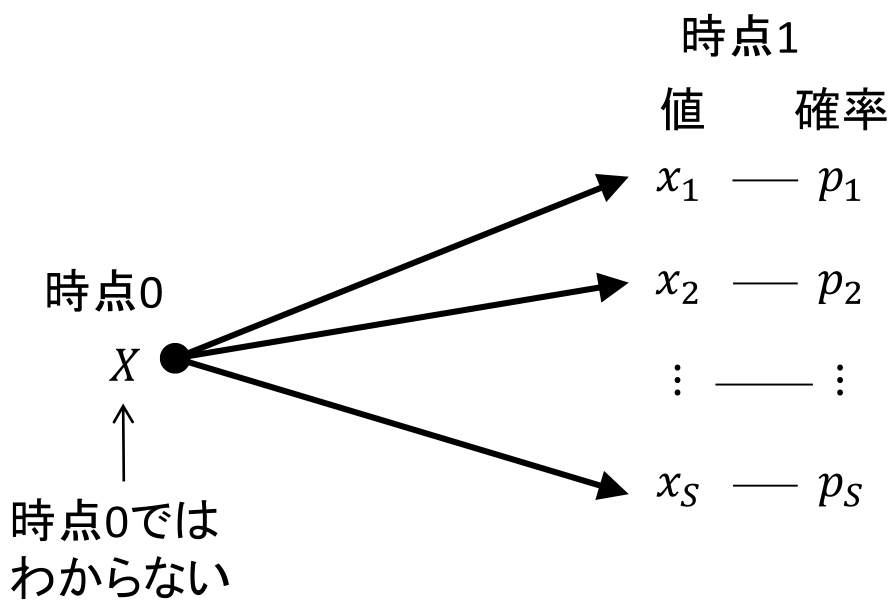


15:50:51

将来に対する確率変数

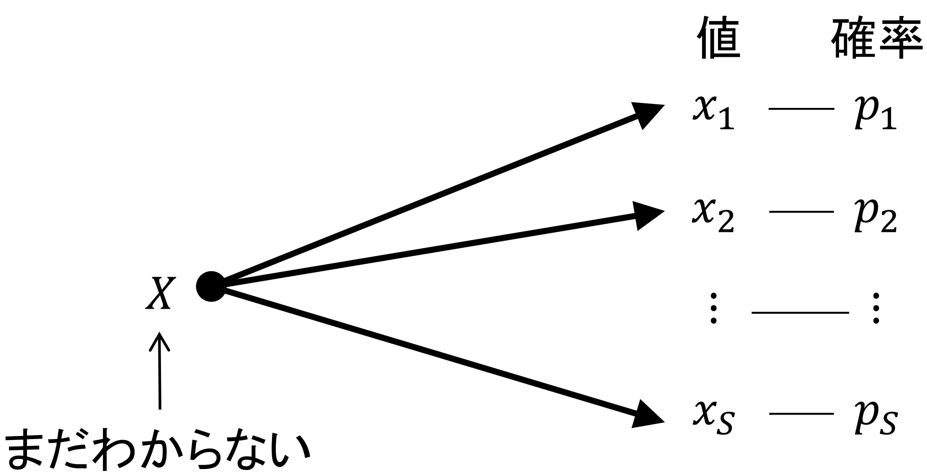
1. 時点0 では確率変数Xはどのような値をとるのかはわからない
   - 時点0 でわかっているのは、時点1 では確率変数Xが$(x_1, x_2, \cdots, x_S)$の内どれかをとること
   - $(x_1, x_2, \cdots, x_S)$のそれぞれの値が生じる確率$(p_1, p_2, \cdots, p_S)$もわかっている

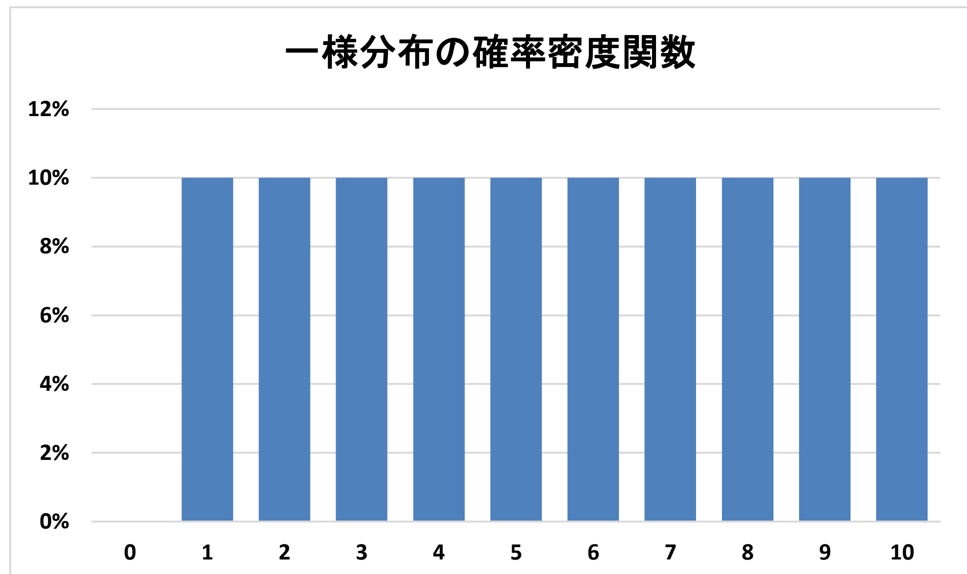
2. 時点1では確率変数Xが$(x_1, x_2, \cdots, x_S)$の内どれかをとる



離散的な確率変数の確率関数

- **各xに対する$f_X(x) = P(\{X = x\})$を確率関数という**
  - $p_1 = P(\{X = x_1\}), p_2 = P(\{X = x_2\}), \ldots, p_S = P(\{X = x_S\})$
  - $P(\{X = x\})$は確率変数Xがあるxとなる確率を意味する
  - 確率関数は確率のリスト
  - 後で連続的な確率変数で紹介される確率密度関数に対応



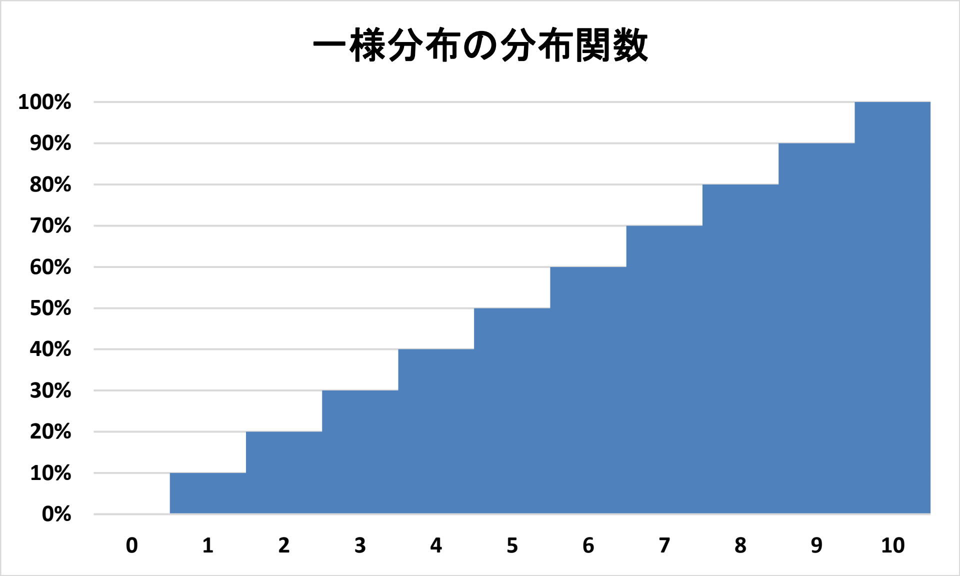


離散的な確率変数の分布関数

- **各xに対する$F_X(x) = P(\{X \le x\})$を累積分布関数（分布関数）という**

$$F_X(x) = P(\{X \le x\}) = P(\{X = x_1\}) + P(\{X = x_2\}) + \cdots + P(\{X = x\})$$

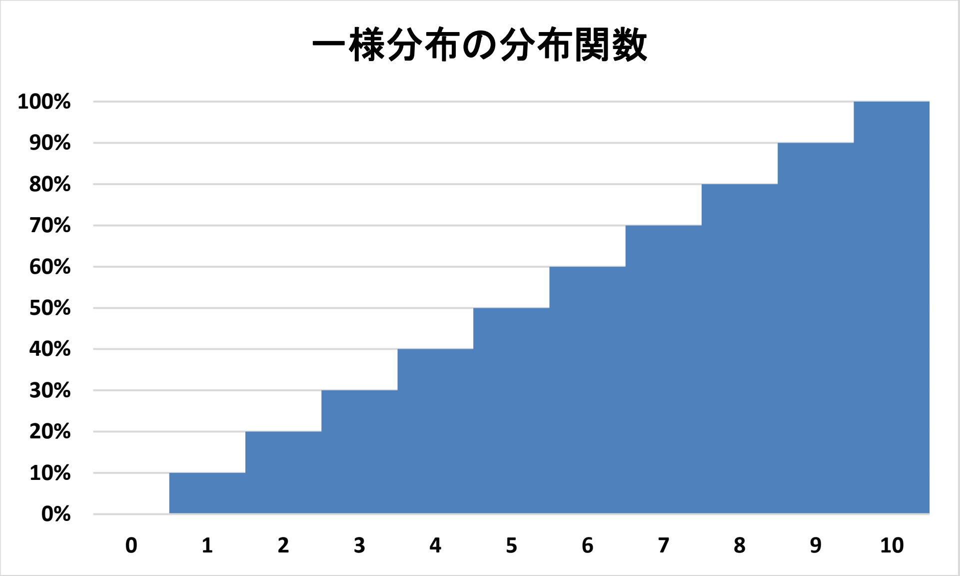
  – 確率関数の和



離散的な確率変数の分布関数

- **各xに対する$F_X(x) = P(\{X \leq x\})$を累積分布関数（分布関数）という**

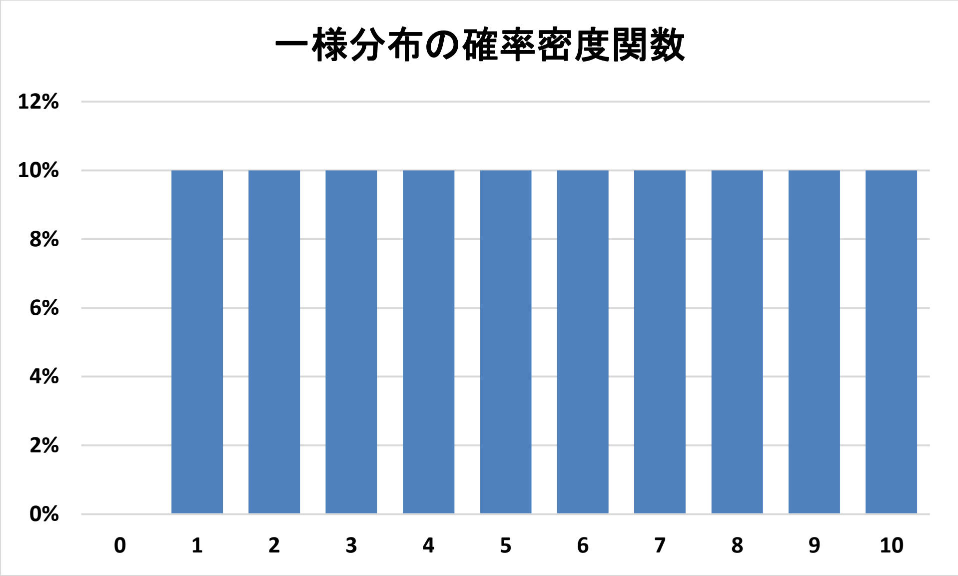
$$F_X(x) = \begin{cases} P(\phi) = 0 & x < 0 \\ P(\{X = 0\}) & 0 \leq x < 1 \\ P(\{X = 0\}) + P(\{X = 1\}) & 1 \leq x < 2 \\ \vdots & \vdots \\ P(\{X = 0\}) + P(\{X = 1\}) + \cdots + P(\{X = S\}) & x \geq S \end{cases}$$

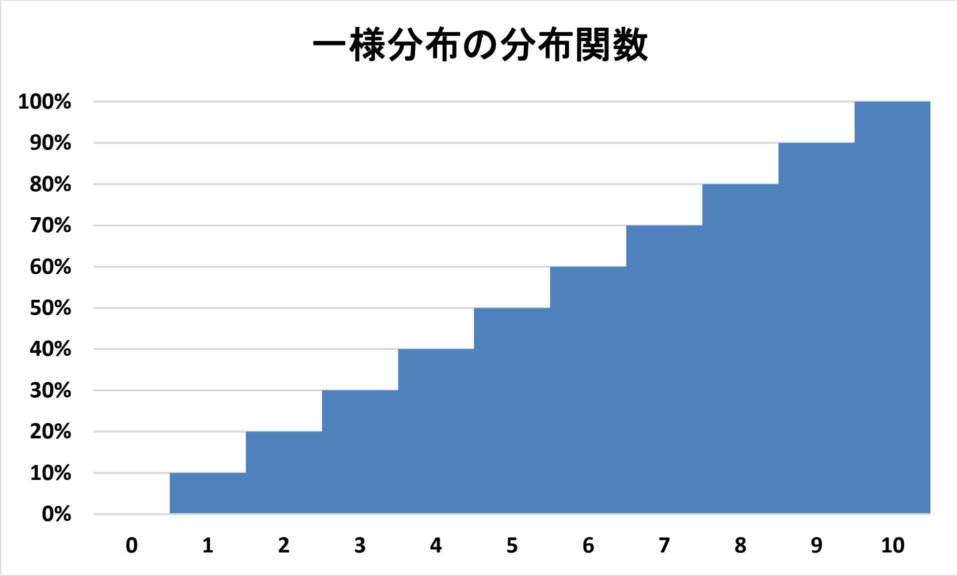


15:50:54

確率密度関数と累積分布関数のイメージ

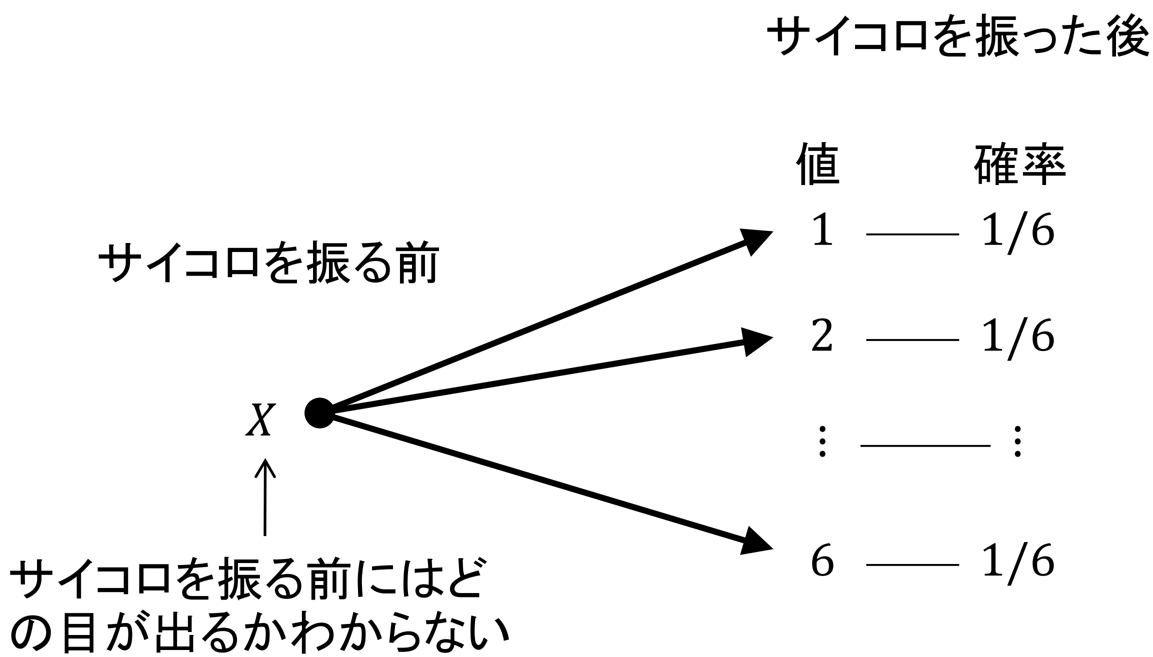
- 確率密度関数はヒストグラムと似ている
  - ヒストグラムと同じように確率密度関数のグラフのx軸は変数の値、y軸は確率
- 累積分布関数はその累積である





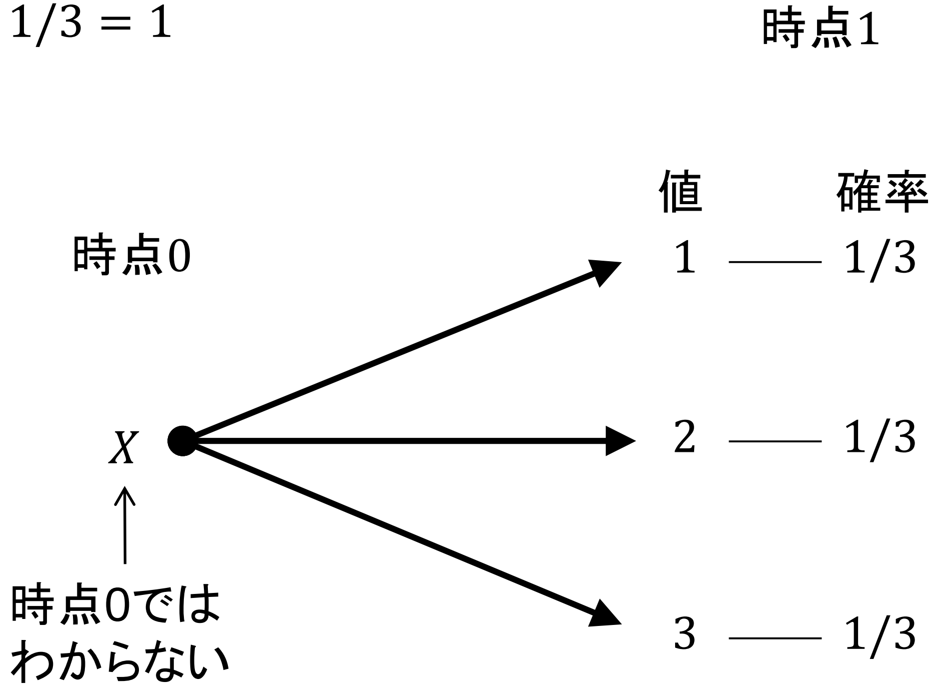
確率変数の例：サイコロ

- サイコロは1から6までの値を取り、
- それぞれの確率は1/6



確率変数: 3つの結果

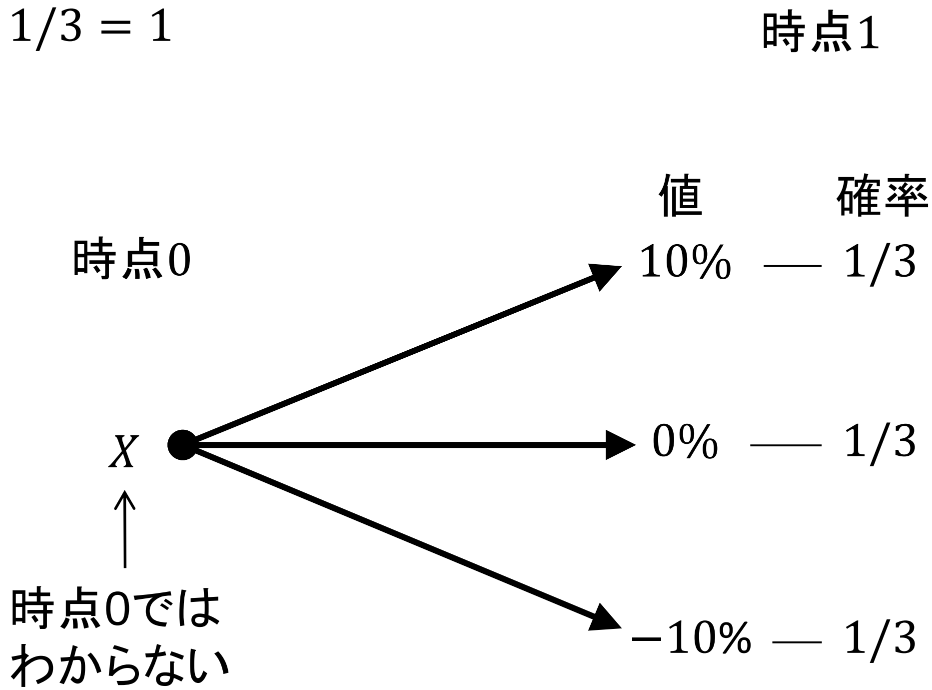
- 確率p_sは
  - $1/3 \geq 0, 1/3 \geq 0, 1/3 \geq 0,$
  - $1/3 + 1/3 + 1/3 = 1$



15:50:57

確率変数: 3つの結果、値を変えた

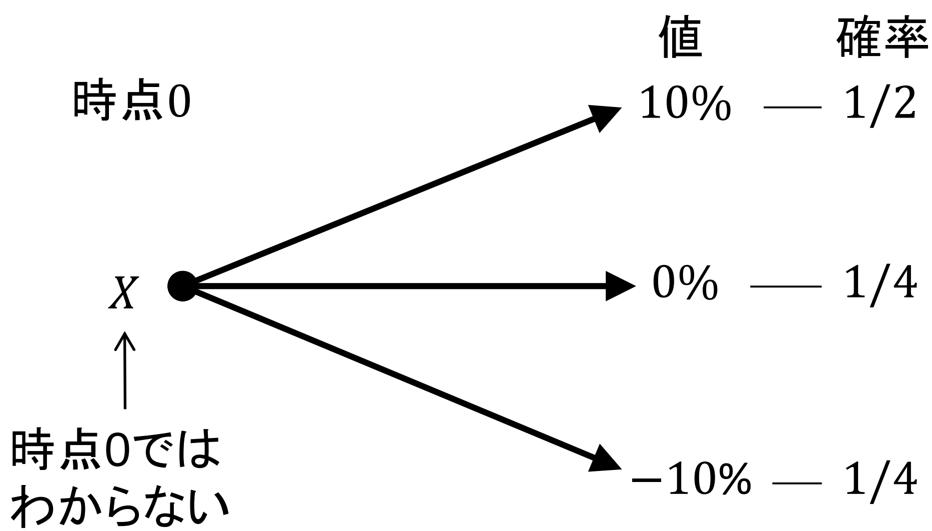
- 確率p_sは
  - $1/3 \geq 0, 1/3 \geq 0, 1/3 \geq 0,$
  - $1/3 + 1/3 + 1/3 = 1$



確率変数: 3つの結果、確率を変えた

- 確率p_sは
  - $1/2 \geq 0, 1/4 \geq 0, 1/4 \geq 0,$
  - $1/2 + 1/4 + 1/4 = 1$

時点1



YOUR TURN

期待値

期待値

- **期待値$E(X)$は**

$$\boldsymbol{E(X) = x_1 \times p_1 + x_2 \times p_2 + \cdots + x_S \times p_S}$$

  - 確率変数Xが取り得る値と各値に対応する確率による加重平均
  - μあるいは$\bar{X}$と表記する場合もある

- 平均とも読み替えられるが、いわゆる算術平均と異なるのは確率$(p_1, p_2, \cdots, p_S)$によって重み付けられている点
  - 算術平均の場合は重みづけはデータの個数の逆数、この場合$1/S$

期待リターン

- 期待リターンとは、リターンの期待値
- 期待リターン$E(r)$は、

$$E(r) = r_1 \times p_1 + r_2 \times p_2 + \cdots + r_S \times p_S$$

  – 確率変数rが取り得る値の各値に対応する確率による加重平均
  – $\bar{r}$と表記する場合もある

期待値の計算のイメージ

- 期待値は、取り得る値とそれに対応する確率を掛け、足し合わせたもの

時点1

期待値は値×確率の合計

$x_1 \times p_1$

$+x_2 \times p_2$

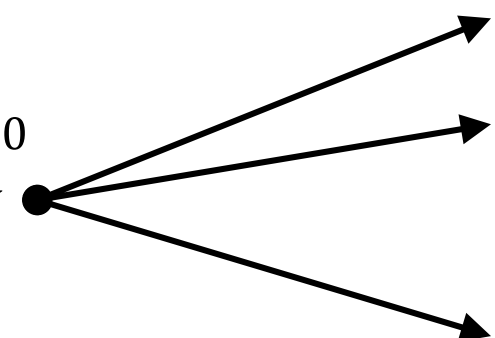
⋮

$+x_S \times p_S$

$= E(X)$

時点0

X



↑

時点0では分からない
ただし、取り得る値$(x_1, x_2, \cdots, x_S)$と対応する確率$(p_1, p_2, \cdots, p_S)$は時点0で分かっている
よって、$E(X)$も分かる

期待値の計算

- 期待値は、取り得る値とそれに対応する確率を掛け、足し合わせたもの

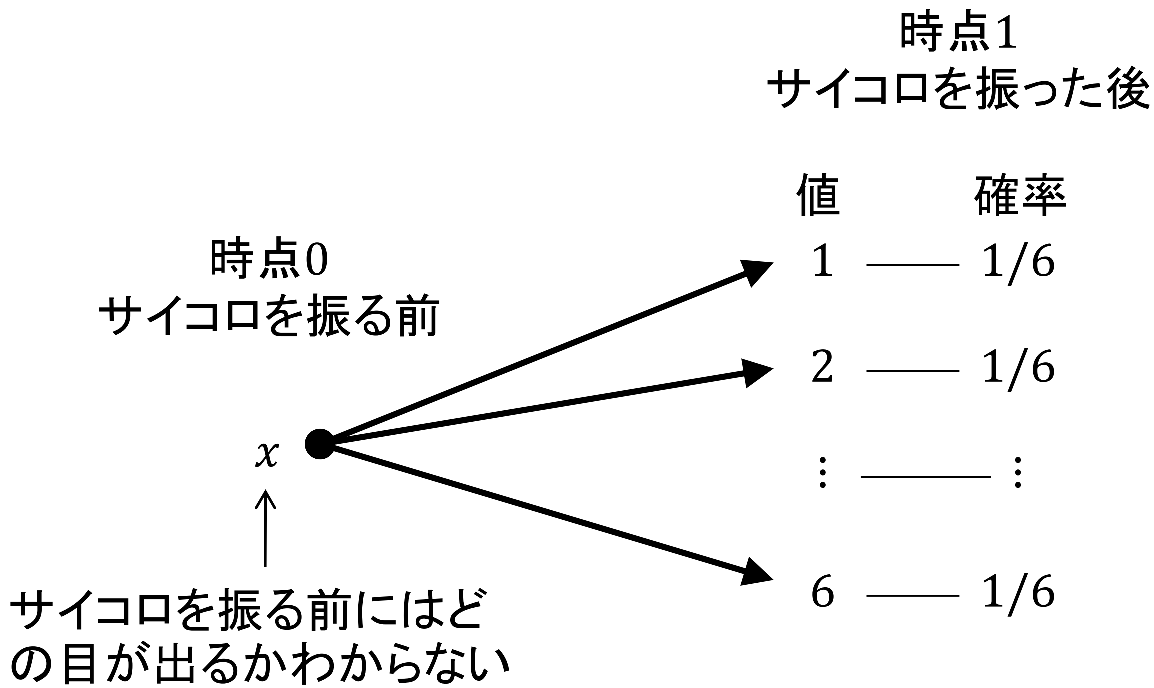
事象	結果	確率	確率で重み付けられた結果
1	x_1	p_1	$x_1 \times p_1$
2	x_2	p_2	$x_2 \times p_2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
S	x_S	p_S	$x_S \times p_S$
		合計	$= E(X)$

期待値の計算例: サイコロの期待値

- サイコロの目の期待値は

$$1\times\frac{1}{6}+2\times\frac{1}{6}+3\times\frac{1}{6}+4\times\frac{1}{6}+5\times\frac{1}{6}+6\times\frac{1}{6}=3.5$$

- 期待値は、実際に取り得る値でなくてもよい
  - *3.5* が期待値だが、これはサイコロの目として取り得る値ではない



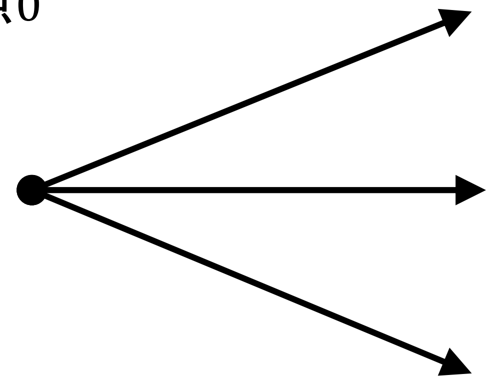
期待値の例

時点1

期待値は値×確率の合計

時点0

x



$10\% \times 1/2$

$0\% \times 1/4$

$-10\% \times 1/4$

$= 0.025 = 2.5\%$

期待値の性質

- もし、c はすでに知っているならば、

$$E(c) = c$$

- もしxが非負、つまり取り得る値がすべて非負ならば、

$$E(X) = x_1 p_1 + \cdots + x_S p_S \geq 0$$

- x, y が確率変数で、a, b が定数ならば、期待値の線形性が成り立つ

$$E(aX + bY) = \Sigma_{s=1}^{S} a x_s p_s + b y_s p_s = aE(X) + bE(Y)$$

YOUR TURN

ばらつきを表す指標：
分散と標準偏差

ばらつきを表す指標：分散と標準偏差

- 確率変数のばらつきを示す尺度を導入する
  - そして、経済学や金融ではこれらをリスクと呼ぶ

ばらつきを表す指標：分散

- **分散$\boldsymbol{Var(X)}$は**

$$\boldsymbol{Var(X) = \left(x_1 - E(X)\right)^2 \times p_1 + \left(x_2 - E(X)\right)^2 \times p_2 + \cdots + \left(x_S - E(X)\right)^2 \times p_S}$$

  - 取り得る値と期待値（平均）の差の平方の期待値
  - 期待値からの乖離を表したもの
  - σ^2やσ_X^2などと表記する場合もある

- 分散を計算するには
  1. 期待値$E(X)$を計算し、
  2. その後に$E\left[\left(X - E(X)\right)^2\right]$を計算する

分散の計算のイメージ

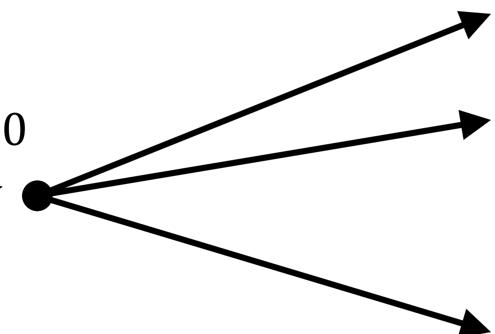
- 分散は、取り得る値の期待値（平均）からの乖離の平方とそれに対応する確率を掛けて、足し合わせたもの

時点1

分散は値と期待値の差の平方の加重平均

時点0

X



$$(x_1 - E[X])^2 \times p_1$$
$$+(x_2 - E[X])^2 \times p_2$$
$$\vdots$$
$$+(x_S - E[X])^2 \times p_S$$
$$= Var(X)$$

時点0では分からない

ただし、取り得る値（$x_1, x_2, \cdots, x_S$）と対応する確率（$p_1, p_2, \cdots, p_S$）は時点0で分かっている

よって、$E(X)$も分かり、$Var(X)$も分かる

分散の計算

- 分散は、取り得る値の期待値（平均）からの乖離の平方とそれに対応する確率を掛けて、足し合わせたもの

事象	結果	確率	期待値からの乖離	期待値からの乖離の平方	確率で重みづけられた期待値からの乖離の平方
1	x_1	p_1	$x_1 - E(X)$	$\left(x_1 - E(X)\right)^2$	$\left(x_1 - E(X)\right)^2 \times p_1$
2	x_2	p_2	$x_2 - E(X)$	$\left(x_2 - E(X)\right)^2$	$\left(x_2 - E(X)\right)^2 \times p_2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
S	x_S	p_S	$x_S - E(X)$	$\left(x_S - E(X)\right)^2$	$\left(x_S - E(X)\right)^2 \times p_S$
				合計	$= Var(X)$

分散の性質

- 分散の性質として以下のようなものがある
  - 計算が簡便的になる場合がある

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

分散の計算例：サイコロ

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

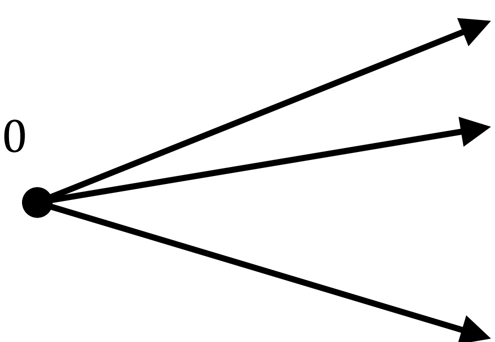
$$= 1^2 \times \frac{1}{6} + 2^2 \times \frac{1}{6} + 3^2 \times \frac{1}{6} + 4^2 \times \frac{1}{6} + 5^2 \times \frac{1}{6} + 6^2 \times \frac{1}{6} - (3.5)^2 = 2.92$$

時点1

分散は値と期待値の差の平方の加重平均

時点0

x



$$(1-3.5)^2 \times 1/6$$
$$+(2-3.5)^2 \times 1/6$$
$$\vdots$$
$$+(6-3.5)^2 \times 1/6$$

$$= 2.92$$

$\uparrow$

時点0では分からない

ただし、取り得る値$(x_1, x_2, \cdots, x_S)$と対応する確率$(p_1, p_2, \cdots, p_S)$は時点0で分かっている

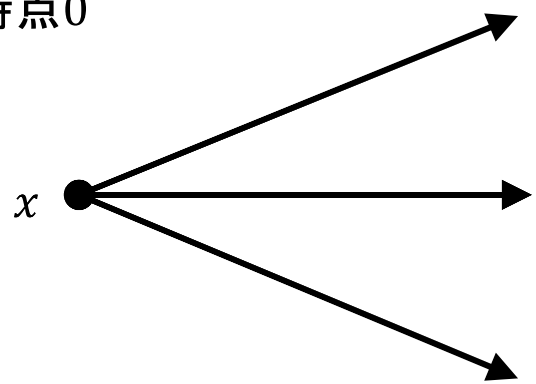
よって、$E[x]$も分かり、$Var[x]$も分かる

分散の計算例

時点0

時点1

分散は値と期待値の差の平方の加重平均



$(10\% - 2.5\%)^2 \times 1/2$

$(0\% - 2.5\%)^2 \times 1/4$

$(-10\% - 2.5\%)^2 \times 1/4$

$= 0.006875$

ばらつきを表す指標：標準偏差

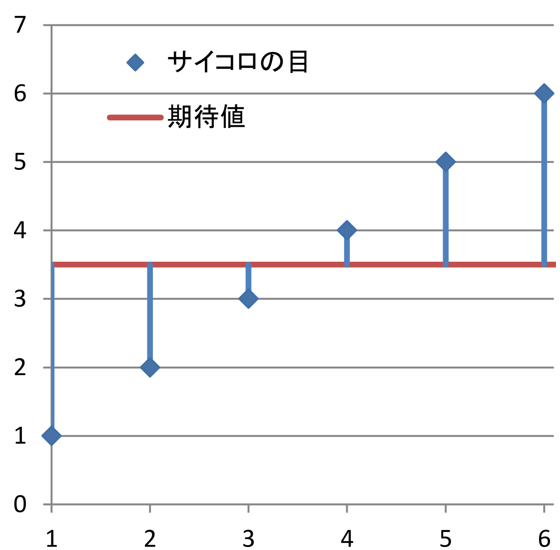
- **標準偏差$\boldsymbol{\sigma_X}$は**
  $\boldsymbol{\sigma_X = \sqrt{Var(X)}}$
  $$= \sqrt{\left(x_1 - E(X)\right)^2 \times p_1 + \left(x_2 - E(X)\right)^2 \times p_2 + \cdots + \left(x_S - E(X)\right)^2 \times p_S}$$
  - 分散の二乗根をとるだけ
  - 期待値からの乖離を表したもの。分散と同じ意味を持つ
  - 標準偏差は期待値と同じ水準で比べてもよく、ファイナンスでは標準偏差は%表示で示すことが多い
    - 分散は期待値の乖離からの平方を取って計算しており、
    - 標準偏差はその分散に対して平方根を取っているため、標準偏差は期待値と同じ水準で比べてもよいということが特徴として挙げられる
  - ファイナンスでは、標準偏差をボラティリティとも呼ぶ
    - 通常、ファイナンスにおけるリスクは標準偏差を指す
    - σなどと表記する場合もある

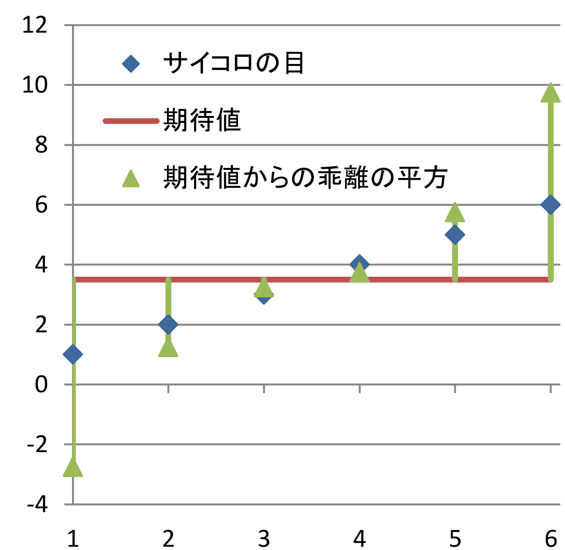
標準偏差の計算例：サイコロ

$$\sigma = \sqrt{Var(X)} = \sqrt{2.92} \approx 1.71$$

分散と標準偏差の計算ステップとイメージ：サイコロ

1. 取り得る値の期待値（平均）を計算（赤い線）
2. 各値と期待値の差を計算（青い線）
3. その差を平方を計算（緑線）
4. 平方した差の平均を計算する。これにより分散を計算
   – 緑の線の長さの平均を計算すると分散
5. 標準偏差の場合は、さらにその平方根を計算





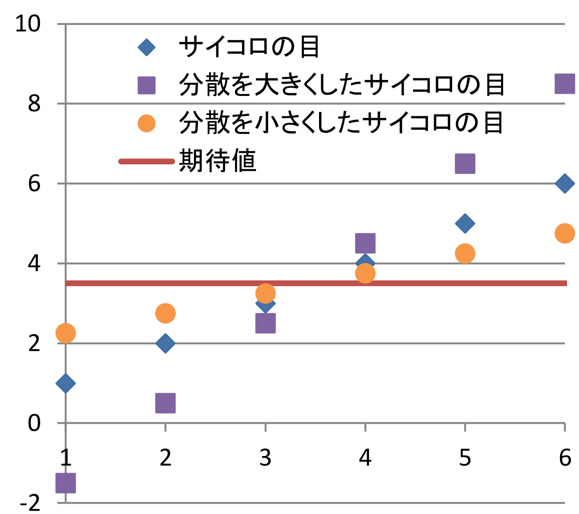
15:50:59

分散の大小のイメージ：サイコロ

- サイコロの目を書き換えてみる
  - 表は、行ごとにサイコロの目を示す

通常のサイコロの目	1	2	3	4	5	6
人工的に分散を大きくしたサイコロの目	-1.5	0.5	2.5	4.5	6.5	8.5
人工的に分散を小さくしたサイコロの目	2.25	2.75	3.25	3.75	4.25	4.75

- 分散を大きくしたサイコロの方が、他のよりもデータとして広がっている
  - ばらつき具合を表し、分散が大きい

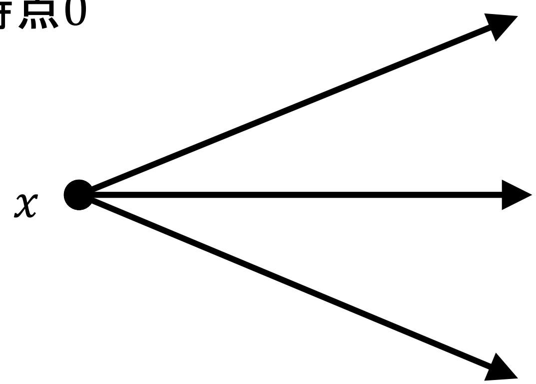


標準偏差の計算例

時点1

分散は値と期待値の差の平方の加重平均

時点0



$$(10\% - 2.5\%)^2 \times 1/4$$

$$(0\% - 2.5\%)^2 \times 1/4$$

$$(-10\% - 2.5\%)^2 \times 1/2$$

$$= 0.006875$$

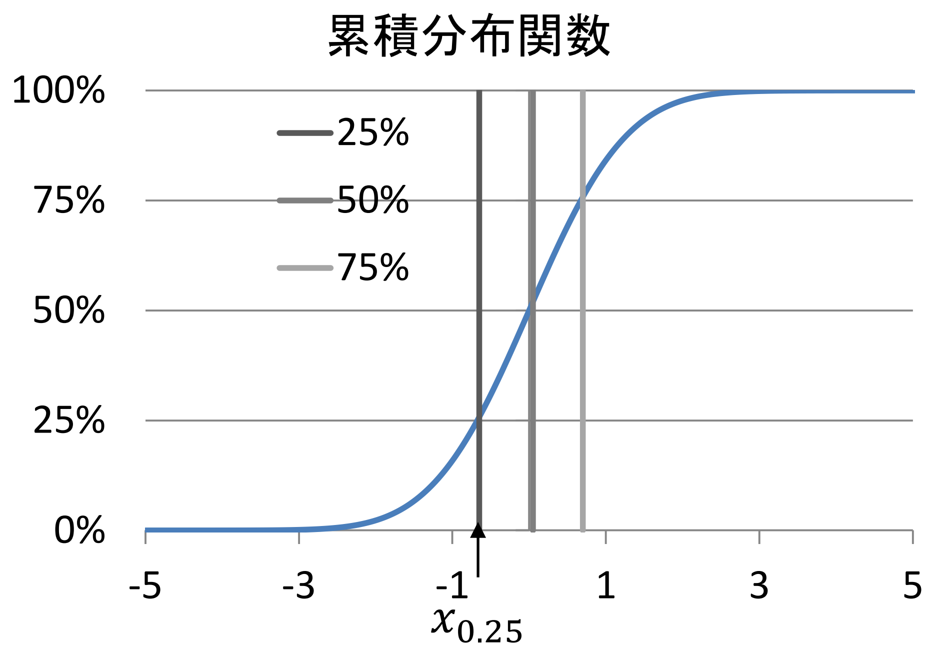
よって、標準偏差は$\sqrt{0.006875} \approx 0.082916$

分散と標準偏差のまとめ

- 分散、標準偏差はデータのばらつきを表している
  - 具体的には、期待値からの乖離がばらつき具合を示す
- 経済学や金融でリスクというとこれらの指標を表す
  - ただし、金融ではリスクというと標準偏差を意味することが多い

確率分布の分位点

- $100q\%$分位点は
  $F_X(x_q) = P(\{X \leq x_q\}) = P(\{X = x_1\}) + \cdots + P(\{X = x_q\}) = q, 0 \leq q \leq 1$
  を満たす点x_qのこと
  - 特に50%分位点をメディアンという
  - $100q\%$分位点は確率をqと$1-q$とに分割する点に他ならない
    $1 - F_X(x_q) = P(\{X = x_{s+1}\}) + \cdots + P(\{X = x_S\}) = 1 - q$



15:51:04

YOUR TURN

連続的な確率変数

連続的な確率変数

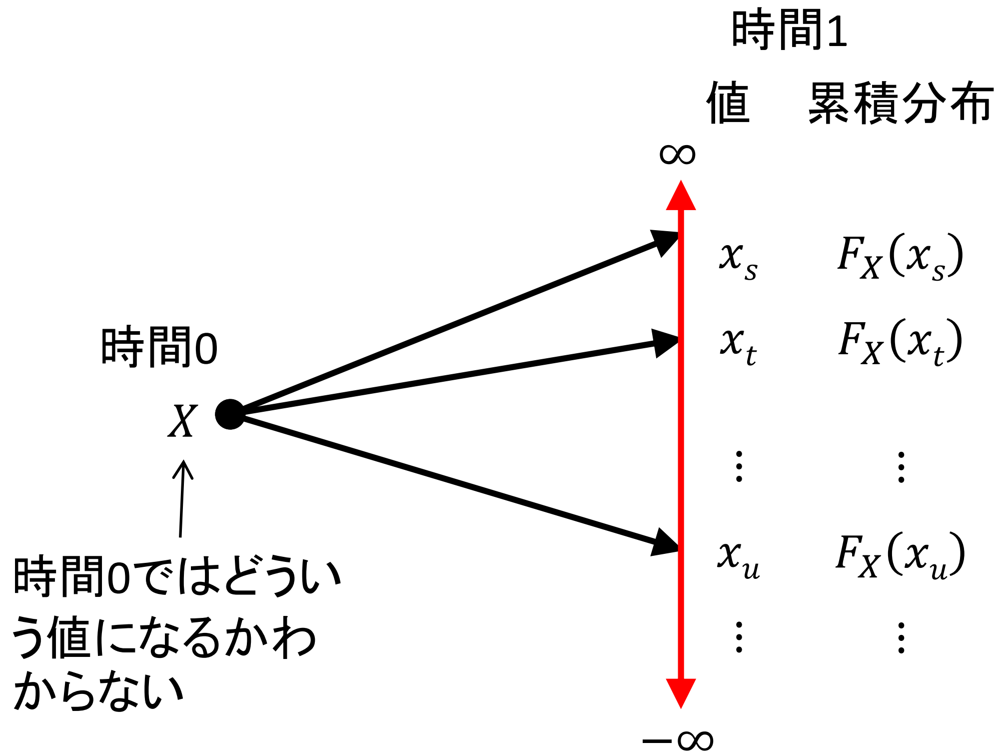
- **連続的な確率変数Xは**
  - **実数$(-\infty, \infty)$の中から取り得る変数であり，**
  - **その事象に対して確率が計算できる**
    - ➢ 離散的な確率変数の場合では、サイコロが1を取ったら、その確率が計算できると同じこと

連続的な確率変数の分布関数

- 各実数値xにはそれぞれの値が発生する**累積分布関数**

$$\boldsymbol{F_X(x) = P(\{X \leq x\})}$$

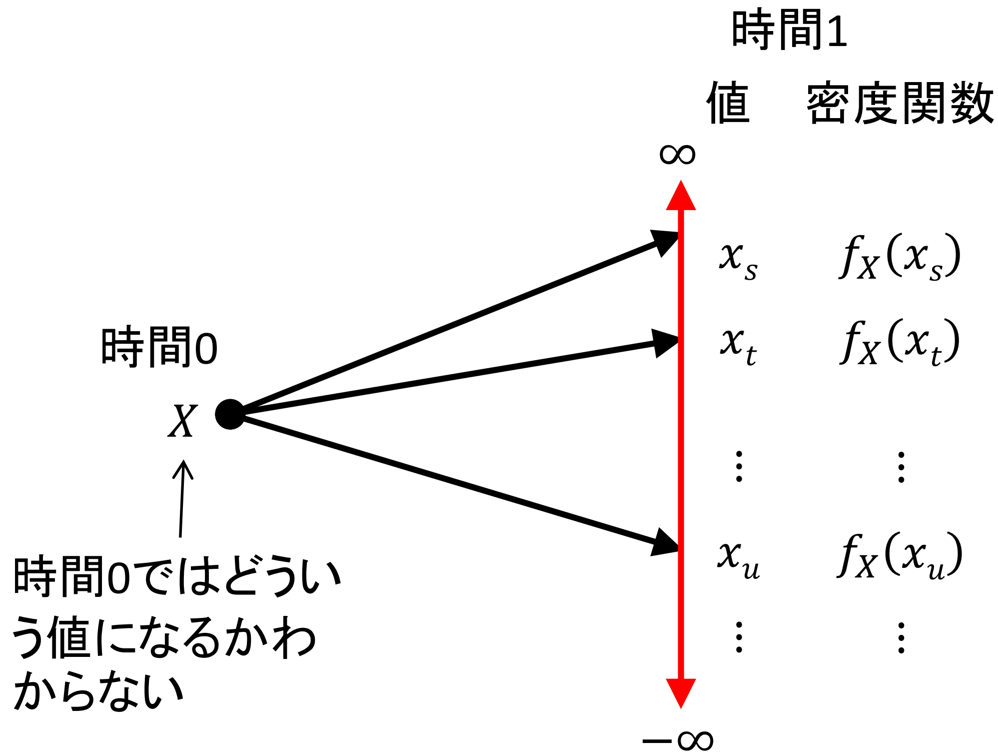
  が対応（分布関数ともいう）
  - $P(\{X \leq x\})$は確率変数Xが実際の値x以下を取る確率のこと
  - 累積分布は$0 \leq F_X(x) \leq 1, F_X(-\infty) = 0, F_X(\infty) = 1$を満たす



15:51:07

連続的な確率変数の密度関数

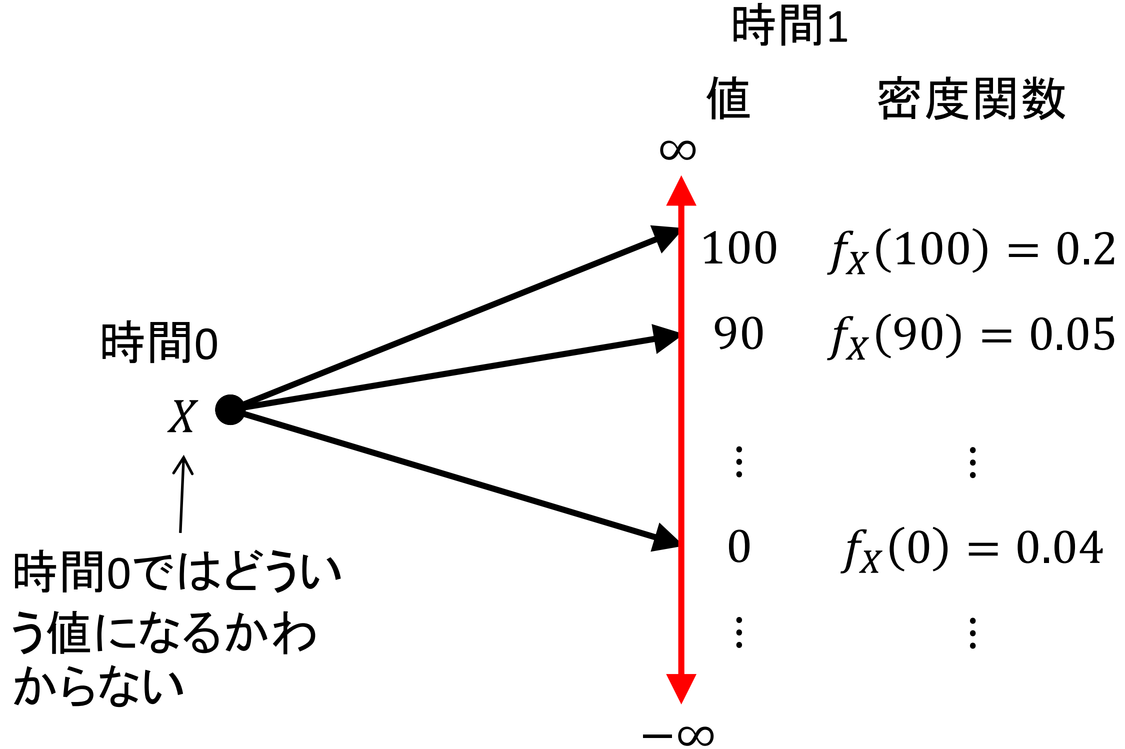
- **任意のxに対して$F_X(x) = \int_{-\infty}^{x} f_X(y)dy$を満たす非負の$f_X(x)(\geq 0)$を確率密度関数という**
  - 密度関数ともいう
  - 各実数値xにはそれぞれの値が発生する確率密度$f_X(x)$が対応



15:51:07

連続的な確率変数のイメージ

- 連続的な（実数値）確率変数xの例は以降に挙げる



確率と密度関数の関係

- $f_X(x) = \lim_{\Delta x \to 0} \frac{P(\{x \leq X \leq x + \Delta x\})}{\Delta x}$
  – 極限を取ると、確率Pから密度関数$f_X(x)$が計算できる
  – あるいはΔxを微小なdxにまで小さくすると、
$$P(\{x \leq X \leq x + dx\}) = f_X(x)dx$$

- $P(\{a < X \leq b\}) = \int_a^b f_X(x)dx$
  – 積分すると、密度関数$f_X(x)$から確率Pが計算できる

分布関数と密度関数の関係

$$F_X(x) = \int_{-\infty}^{x} f_X(y)dy \quad \Leftrightarrow \quad \frac{dF_X(x)}{dx} = f_X(x)$$

- 分布関数から密度関数が計算でき、
- 密度関数から分布関数が計算できる
- 微分積分学の基本定理

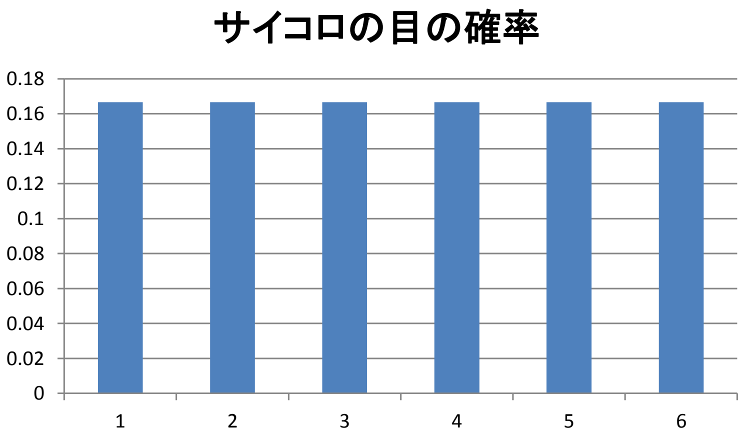
- $P(\{a < X \leq b\}) = F_X(b) - F_X(a) = \int_a^b f_X(x)dx$
  - 確率と分布関数と密度関数は
    - ➢ まとめているか瞬間的に見ているか、
    - ➢ まとめている場合はどのようにまとめているか

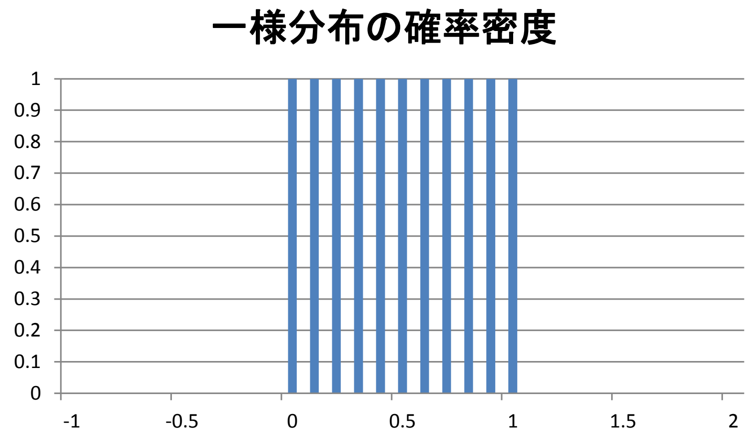
  の違い

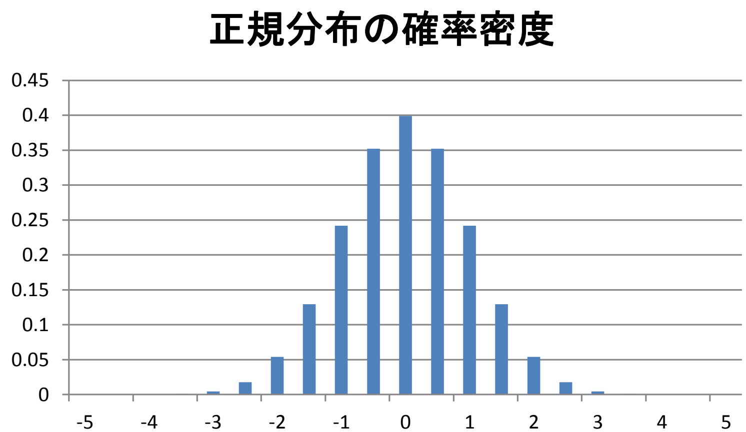
確率と密度関数のイメージ

- 密度関数は確率ではないが，

  確率 ≈ 確率密度 × 確率変数が取る区間の幅
- 確率あるいは確率密度のグラフがヒストグラムに似ているのは偶然ではない

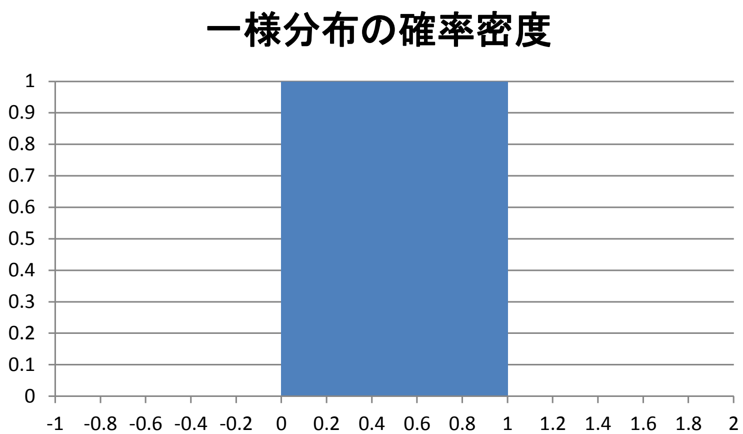


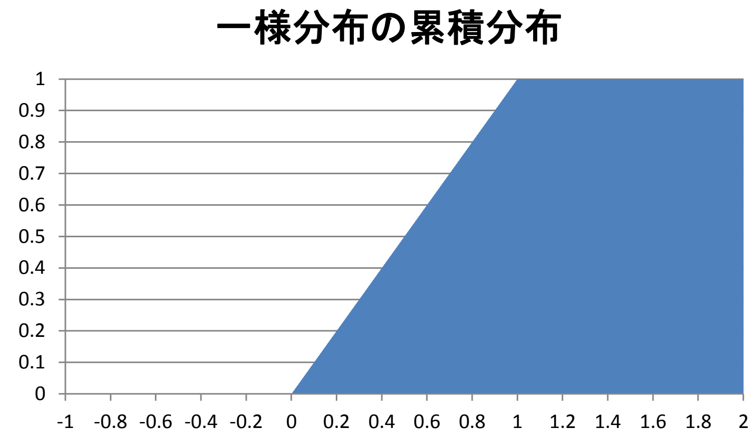


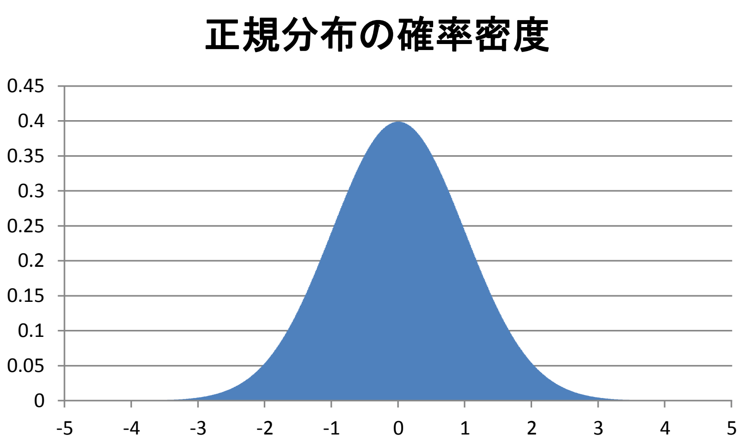


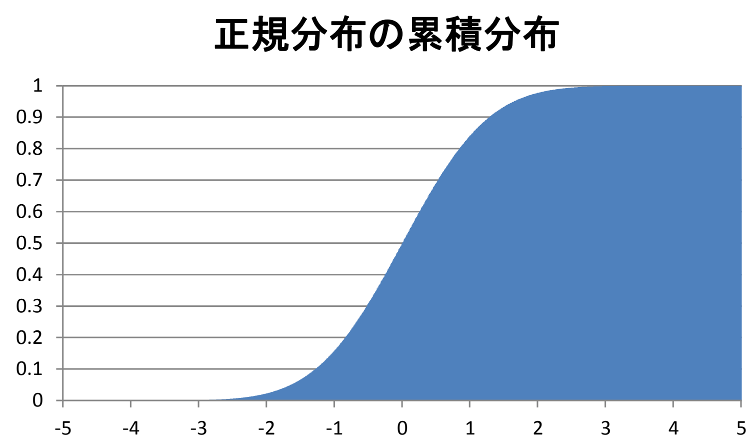
密度関数と分布関数のイメージ

- 分布関数は密度関数による面積
  - $\int_{-\infty}^{x} f_X(y)dy$: ある点yにおいてそれまでの確率密度を足し合わせたもの



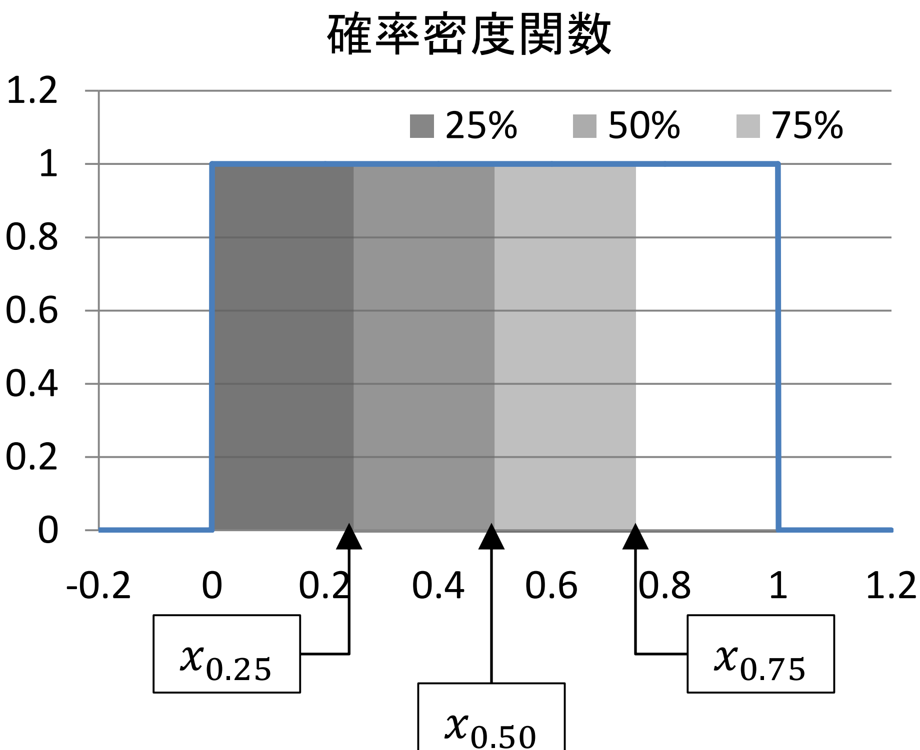


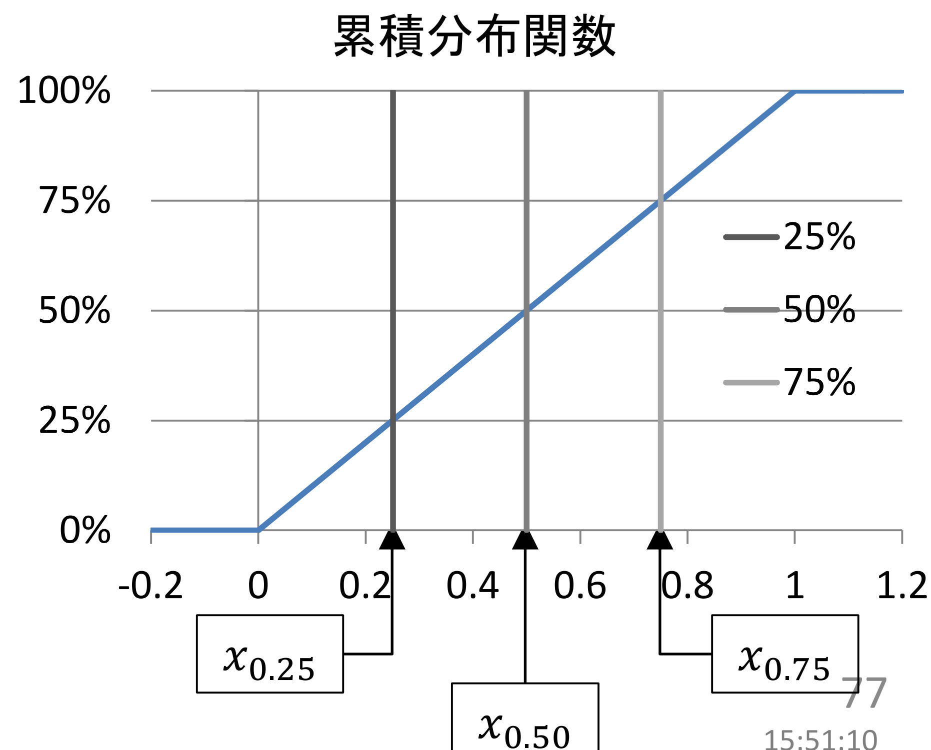




一様分布の密度関数と分布関数のイメージ

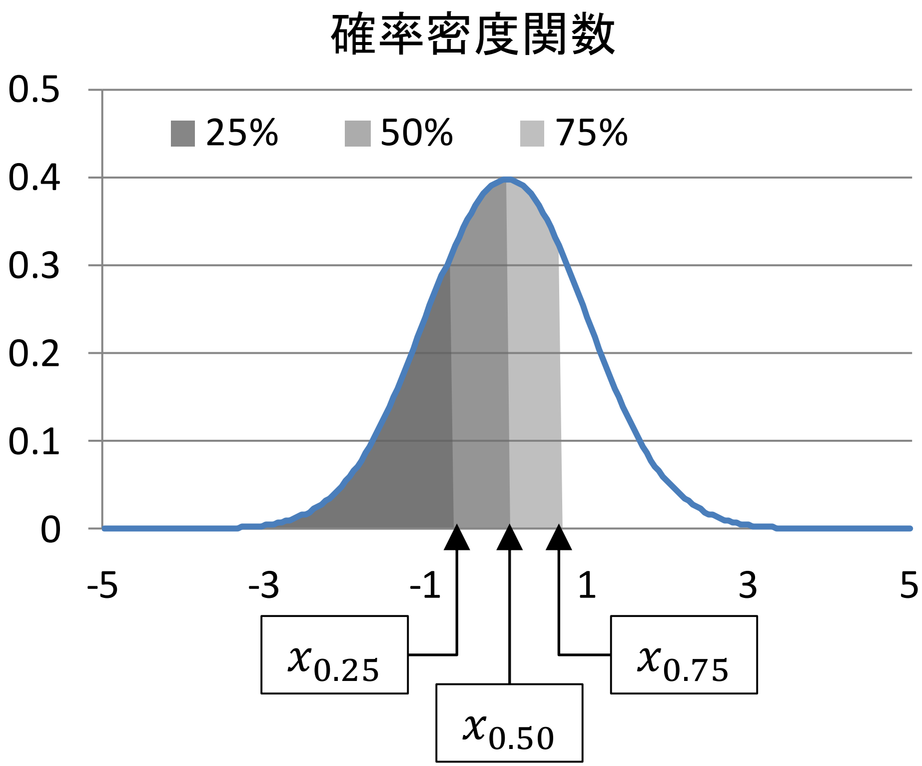
- 密度関数の面積が確率
- xとxまでの面積を計算した確率を，改めてプロットしたのが分布関数
  - 25%と$x_{0.25}$（確率25%に対するx）
  - 50%と$x_{0.50}$（確率50%に対するx）
  - 75%と$x_{0.75}$（確率75%に対するx）

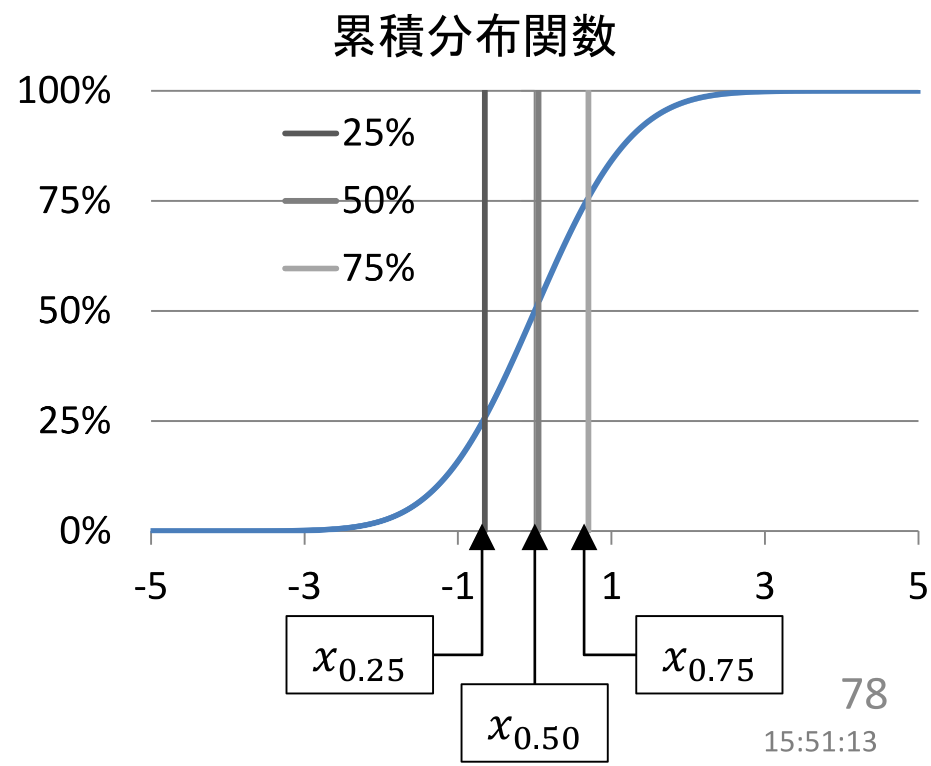




正規分布の密度と分布関数のイメージ

- 密度関数の面積が確率
- xとxまでの面積を計算した確率を，改めてプロットしたのが分布関数
  - 25%と$x_{0.25}$（確率25%に対するx）
  - 50%と$x_{0.50}$（確率50%に対するx）
  - 75%と$x_{0.75}$（確率75%に対するx）





YOUR TURN

期待値：連続的な確率変数の場合

- **期待値$E(X)$は**

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

  - 確率変数Xが取り得る値の各値に対応する確率による加重平均
  - μあるいは$\bar{X}$と表記する場合もある

- 平均とも読み替えられるが、密度関数$f_X(x)$によって重み付けられている

分散:連続的な確率変数の場合

- **分散$\boldsymbol{Var(X)}$は**

$$\boldsymbol{Var(X) = \int_{-\infty}^{\infty} \left(x - E(X)\right)^2 f_X(x)dx = E\left[\{X - E(X)\}^2\right]}$$

  - 取り得る値と期待値(平均)の差の平方の期待値
  - 期待値からの乖離を表したもの
  - σ^2やσ_X^2などと表記する場合もある

- 分散を計算するには
  1. 期待値$E(X)$を計算し、
  2. その後に$E\left[\left(X - E(X)\right)^2\right]$を計算する

標準偏差：連続的な確率変数の場合

- **標準偏差σ_Xは**

$$\sigma_X = \sqrt{Var(X)}$$
$$= \sqrt{\int_{-\infty}^{\infty} \left(x - E(X)\right)^2 f_X(x)dx}$$

  - 分散の2乗根をとるだけ
  - 期待値からの乖離を表したもの。分散と同じ意味を持つ
  - σなどと表記する場合もある

期待値の性質

- もし、c はすでに知っているならば、
$$E(c) = c$$

- もしxが非負、つまり取り得る値がすべて非負ならば、
$$E(X) \geq 0$$

- x, y が確率変数で、a, b が定数ならば、期待値の線形性が成り立つ
$$E(aX + bY) = aE(x) + bE(y)$$

確率分布の分位点

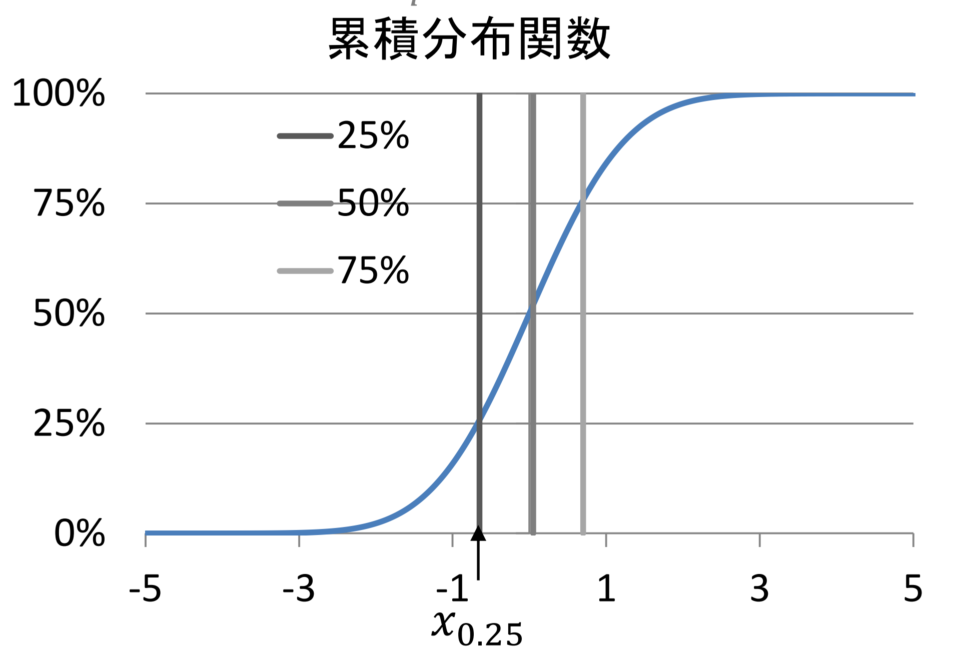
- $100q\%$分位点は

$$F_X(x_q) = \int_{-\infty}^{x_q} f_X(y)dy = q, 0 \leq q \leq 1$$

を満たす点x_qのこと

  - 特に50%分位点をメディアンという
  - $100q\%$分位点は確率をqと$1-q$とに分割する点に他ならない

$$1 - F_X(x_q) = \int_{x_q}^{\infty} f_X(y)dy = 1 - q$$



YOUR TURN