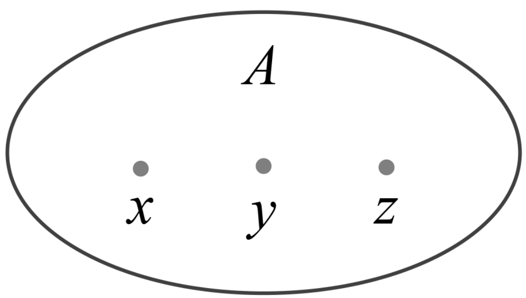
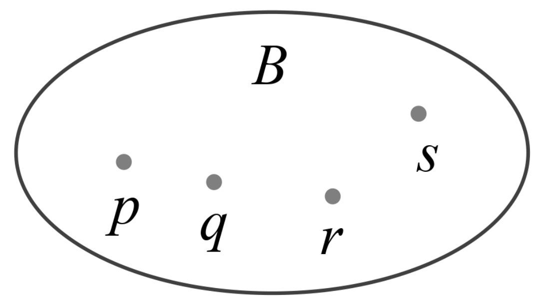
1-6-1　集合と順列

集合の概念

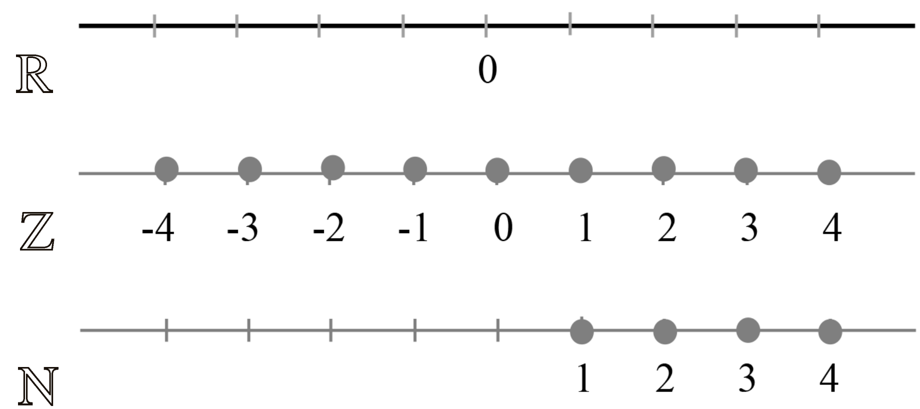
- 集合とは，ある特定の「もの」の集まり のことです.
- 集合 A を構成する「もの」のことを 要素または元(げん)とよびます.
  - x が A の要素であるとき，$x \in A$ または $A \ni x$ と書きます.
  - p が A の要素でないとき，$p \notin A$ または $A \not\ni p$ と書きます.
- 有限個の要素からなる集合を有限集合とよび，
  無限個の要素からなる集合を無限集合とよびます.





基本的な数の集合

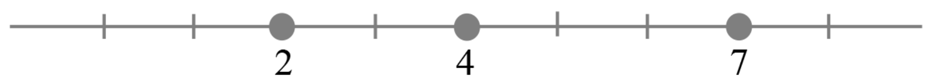
- 実数の集合 $\mathbb{R}$.
- 整数の集合 $\mathbb{Z}$.
- 自然数の集合 $\mathbb{N}$.
- 有理数の集合 $\mathbb{Q}$.



- これらは，すべて 無限集合です.
- 例： $16 \in \mathbb{N}$, $\quad 24.2 \notin \mathbb{N}$, $\quad 24.2 \in \mathbb{Q}$, $\quad \sqrt{5} \notin \mathbb{Q}$, $\quad \sqrt{5} \in \mathbb{R}$.

集合の表し方 (1/2)

- 例として，$\{2, 4, 7\}$ は,「3 つの実数 2, 4, 7 からなる集合」のことです.
  - 要素を書き並べる順番を変えても，集合としては同じものです.
    - つまり，$\{2, 4, 7\} = \{2, 7, 4\} = \{7, 4, 2\} = \{4, 7, 2\} = \cdots$.



- $\{5, 6, 7, \ldots, 30\}$ は,「5 以上 30 以下の自然数からなる集合」のことです.
- $\{1, 2, 3, \ldots\} = \mathbb{N}$.
- $\{0, \pm 1, \pm 2, \pm 3, \ldots\} = \mathbb{Z}$.

集合の表し方 (2/2)

- 「5 以上 30 以下の自然数からなる集合」は，$\{5, 6, 7, \ldots, 30\}$ と表しました.
  - これとは別の表し方として，

$$\{x \mid x \text{ は 5 以上 30 以下の自然数 }\},$$

または，$\{x \in \mathbb{N} \mid 5 \leqq x \leqq 30\}$ もあります.

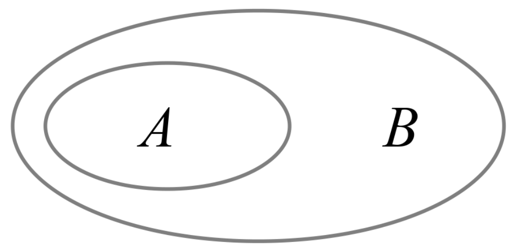
- $\{x \in \mathbb{R} \mid 5 \leqq x \leqq 30\}$ は,「5 以上 30 以下の実数からなる集合」です.
- $\{2i \mid i \in \mathbb{Z}\}$ は，偶数の集合です.
- $\{2i - 1 \mid i \in \mathbb{Z}\}$ は，奇数の集合です.

空集合

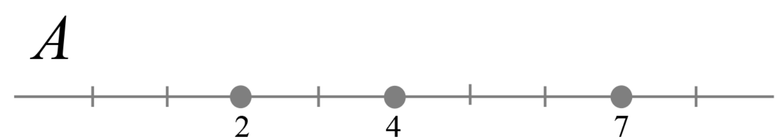
- 要素を 1 つももたない集合を，空集合とよびます.
  - 空集合は，$\emptyset$ や $\{\}$ などで表します.
- 例： $\{x \in \mathbb{R} \mid x^2 = -1\} = \emptyset$.
- 例： $\{x \in \mathbb{N} \mid x \leqq -5\} = \emptyset$.

部分集合

- 集合 A の要素がすべて集合 B の要素でもあるとき，A を B の部分集合とよびます.
  - このとき，$A \subset B$ または $B \supset A$ と書きます.



- $A \subset B$ かつ $A \supset B$ のとき，$A = B$ と書きます.

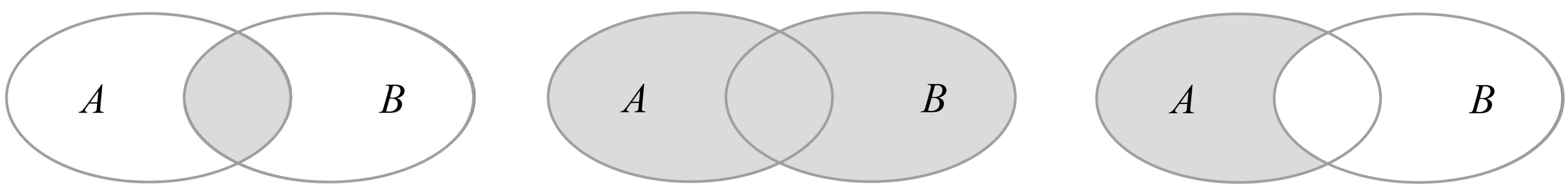


- 例： $A = \{2, 4, 7\}$ のとき，
  - $A \subset \{2, 4, 6, 7\}$,　$A \subset \mathbb{N}$,　$\{2, 4\} \subset A$.
  - $\{2, 5\}$ は，A の部分集合ではありません.
  - A の部分集合は，全部で 8 個あります：

$$\emptyset\,,\ \{2\}\,,\ \{4\}\,,\ \{7\}\,,\ \{2, 4\}\,,\ \{4, 7\}\,,\ \{2, 7\}\,,\ \{2, 4, 7\}\,.$$

共通部分，和集合，差集合 (1/2)

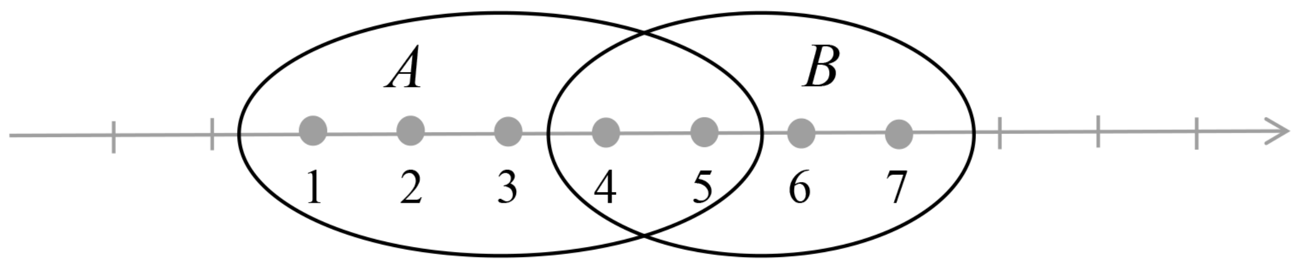
- 2 つの集合 A と B に対して，次の集合が定義できます.
- A と B の共通部分（または，交わり）$A \cap B$ とは，
  A と B の両方に含まれる要素をすべて集めた集合のことです.
- A と B の和集合 $A \cup B$ とは，
  A の要素と B の要素をすべて集めた集合のことです.
- A と B の差集合 $A - B$ とは，
  A に含まれるが B に含まれない要素をすべて集めた集合のことです.



$A \cap B$ $A \cup B$ $A - B$

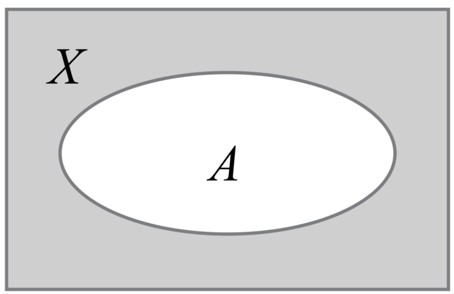
共通部分，和集合，差集合 (2/2)

- 例： $A = \{1, 2, 3, 4, 5\}$, $B = \{4, 5, 6, 7\}$ に対して，
  - $A \cap B = \{4, 5\}$.
  - $A \cup B = \{1, 2, 3, 4, 5, 6, 7, \}$.
  - $A - B = \{1, 2, 3\}$.
  - $B - A = \{6, 7\}$.



補集合

- 補集合
  - 集合 X と，その部分集合 A に対して，
    $X - A$ を A の（X における）補集合とよび，
    A^{c} や $X \setminus A$ などで表します．
  - このとき，X を全体集合とよびます．



（なお，このような図は，ベン図とよばれます．）

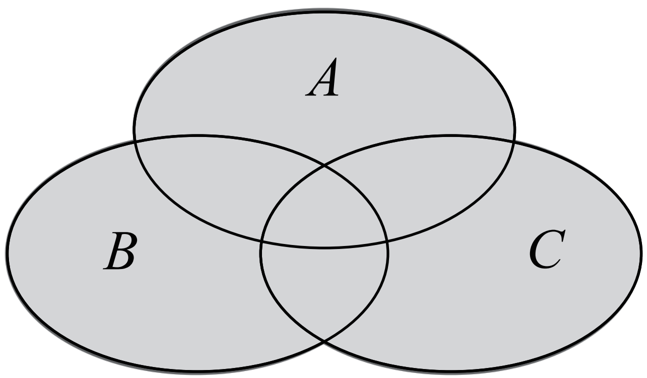
- 例： $X = \{1, 2, 3, 4, 5, 6\}$, $A = \{2, 3, 4\}$ とすると，$A^{c} = \{1, 5, 6\}$ です．
- 例： $X = \mathbb{Z}$, $A = \{x \in \mathbb{Z} \mid x$ は 3 で割ると 1 余る $\}$ とすると，
  A^{c} は「3 の倍数と，3 で割ると 2 余る整数とを，すべて集めた集合」です．

結合法則と分配法則 (1/2)

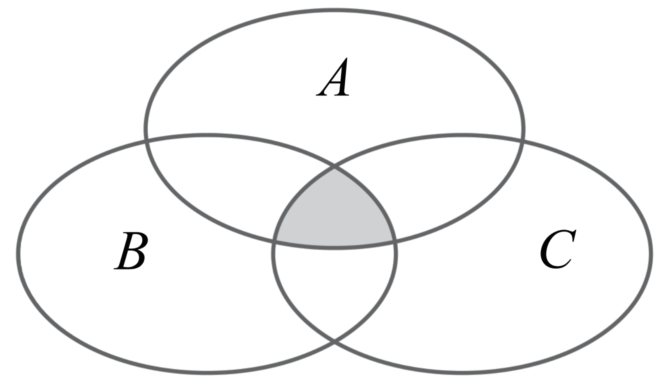
- 結合法則

  (i) $(A \cup B) \cup C = A \cup (B \cup C)$.

  (ii) $(A \cap B) \cap C = A \cap (B \cap C)$.



(i)



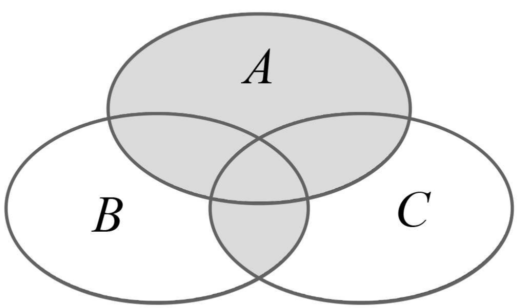
(ii)

結合法則と分配法則 (2/2)

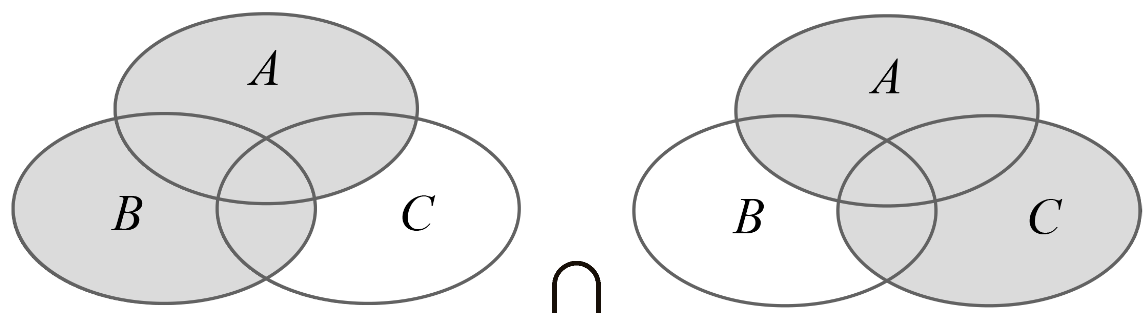
- 分配法則

(i) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

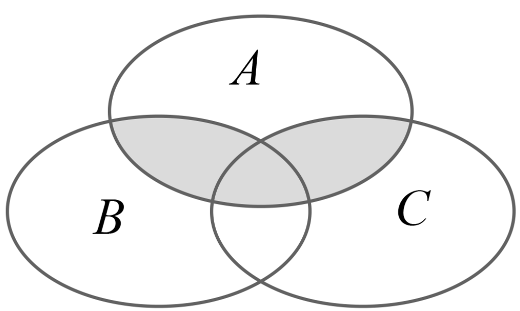
(ii) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.



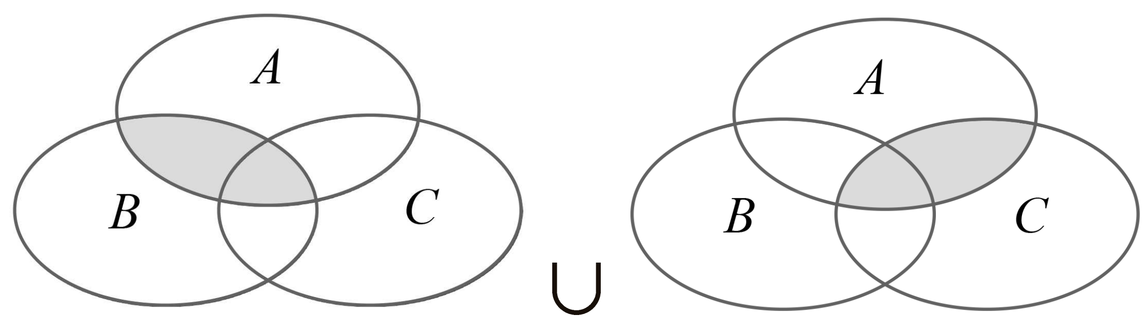
(i) の左辺



(i) の右辺



(ii) の左辺



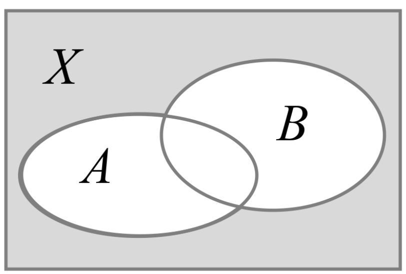
(ii) の右辺

ド・モルガンの法則

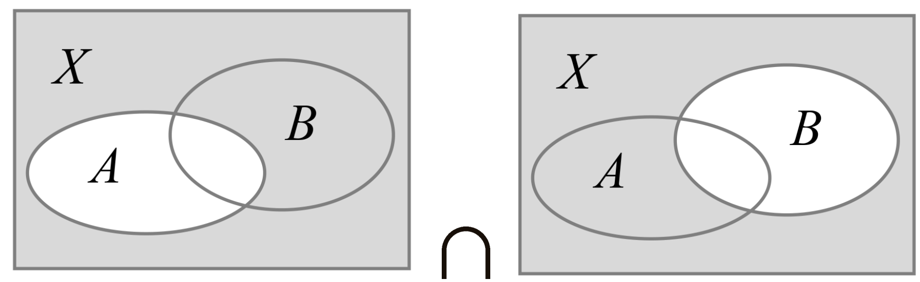
- 集合 X, A, B に対して,

  (i) $X - (A \cup B) = (X - A) \cap (X - B)$.

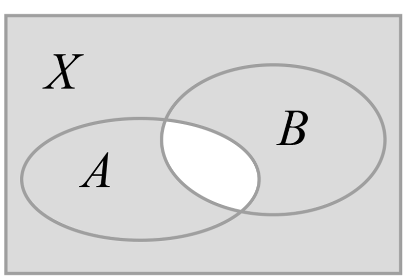
  (ii) $X - (A \cap B) = (X - A) \cup (X - B)$.



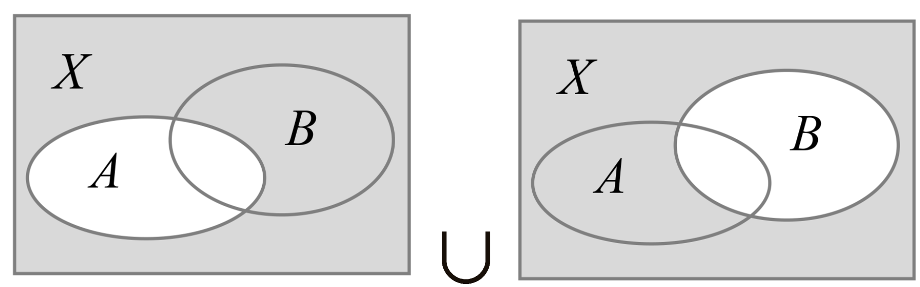
(i) の左辺



(i) の右辺



(ii) の左辺

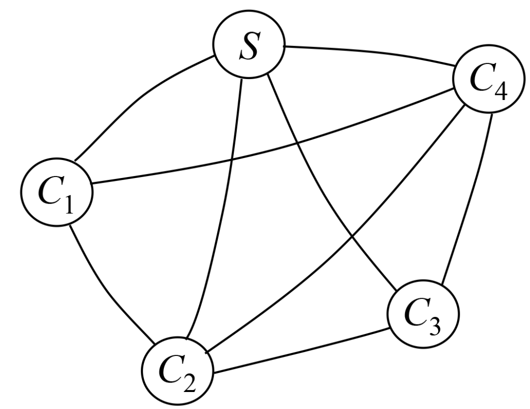


(ii) の右辺

順列：導入

- 例として，巡回セールスマン問題とよばれる問題をとりあげます.

- 都市 s から出発し，都市 $c_1, \ldots, c_4$ のすべてをちょうど 1 回ずつ訪れてから s に戻る経路のうち，移動距離が最小のものを求めて下さい.



- 候補となる経路は，

$$
\begin{array}{l}
s \to c_1 \to c_2 \to c_3 \to c_4 \to s \\
s \to c_1 \to c_2 \to c_4 \to c_3 \to s \\
s \to c_1 \to c_3 \to c_2 \to c_4 \to s \\
s \to c_1 \to c_3 \to c_4 \to c_2 \to s \\
s \to c_1 \to c_4 \to c_2 \to c_3 \to s \\
s \to c_1 \to c_4 \to c_3 \to c_2 \to s \\
s \to c_2 \to \cdots \\
\qquad \vdots
\end{array}
$$

  - 最初を c_1 と決めたら，6 通り.
  - 最初の決め方は 4 通り $\to$ 全部で $4 \cdot 3 \cdot 2 \cdot 1 = 24$ 通り.

順列とは

- いくつかのものを一列に並べるとき，その並べ方の１つ１つのことを，順列とよびます.
- 4 個の都市の順列の総数は $4 \cdot 3 \cdot 2 \cdot 1 = 24$ 通り.
- n 個の（相異なるものの）順列の総数は $n!$ です.
  - $n! = n \cdot (n-1) \cdots 3 \cdot 2 \cdot 1$ のことを，n の階乗とよびます.
    - $0! = 1$ と決めておく.
    - $n!$ は，n が大きくなると急激に大きくなります.
- n 個から r 個を選んで並べる場合は，順列の総数は

$$\underbrace{n \cdot (n-1) \cdots (n-r+2) \cdot (n-r+1)}_{r \text{ 個の積}}.$$

  - この数のことを，${}_n\mathrm{P}_r$ と書きます.
    - ${}_n\mathrm{P}_r = \dfrac{n!}{(n-r)!}$,　${}_n\mathrm{P}_n = n!$ です.

（狭義の）組合せとは

- n 個のものから r 個を選ぶときの，選び方の総数のことを，組合せとよび，${}_n\mathrm{C}_r$ や $\dbinom{n}{r}$ で表します．
  - 順列では順番を区別しますが，組合せでは順番を区別しません．
  - 選んできた r 個の順列の総数は $r!$ ですので，
$$ {}_n\mathrm{C}_r = \frac{{}_n\mathrm{P}_r}{r!} = \frac{n!}{(n-r)!r!} \, . $$
  - ${}_n\mathrm{C}_0 = 1 \, , \quad {}_n\mathrm{C}_n = 1 \, .$
  - ${}_{n+1}\mathrm{C}_r = {}_n\mathrm{C}_{r-1} + {}_n\mathrm{C}_r \, .$　　　（再帰的な定義）
    - $n+1$ 個目を選ぶ場合は，n 個目までからはあと $r-1$ 個を選ぶので，${}_n\mathrm{C}_{r-1}$.
    - $n+1$ 個目を選ばない場合は，n 個目までから r 個すべてを選ぶので，${}_n\mathrm{C}_r$.

2 項定理

- 母関数
  - 数列 $a_0, a_1, a_2, \ldots, a_n$ に対して，関数
  $$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$
  のことを（x を変数とする）母関数とよびます.
- 数列 ${}_n\mathrm{C}_0, {}_n\mathrm{C}_1, {}_n\mathrm{C}_2, \ldots, {}_n\mathrm{C}_n$ の母関数は $(1+x)^n$ です，つまり，
$$(1+x)^n = {}_n\mathrm{C}_0 + {}_n\mathrm{C}_1 x + {}_n\mathrm{C}_2 x^2 + \cdots + {}_n\mathrm{C}_{n-1} x^{n-1} + {}_n\mathrm{C}_n x^n\ .$$
- これを一般化して，
$$(x+y)^n = {}_n\mathrm{C}_0 y^n + {}_n\mathrm{C}_1 x y^{n-1} + {}_n\mathrm{C}_2 x^2 y^{n-2} + \cdots + {}_n\mathrm{C}_{n-1} x^n y + {}_n\mathrm{C}_n x^n\ .$$
  - これを，2 項定理とよびます．また，${}_n\mathrm{C}_r$ を 2 項係数とよびます.
- 例：
  - $(x+y)^6$ を展開したときの x^2y^4 の係数は，${}_6\mathrm{C}_2 = 15$.
  - $(x-3y)^7$ を展開したときの x^4y^3 の係数は，${}_7\mathrm{C}_4 \cdot (-3)^3 = -945$.

2 項定理の応用

- $\displaystyle\sum_{k=0}^{n} {}_n\mathrm{C}_k = 2^n$.
  - $\because\ 2^n = (1+1)^n = \displaystyle\sum_{k=0}^{n} {}_n\mathrm{C}_k 1^k 1^{n-k}$.

    （n 個それぞれについて，選ぶか否かの 2 通りがある，と考えることもできます.）

- ${}_{m+n}\mathrm{C}_r = \displaystyle\sum_{k=0}^{r} {}_m\mathrm{C}_k {}_n\mathrm{C}_{r-k}$.
  - $\because\ (1+x)^{m+n} = (1+x)^m(1+x)^n = \displaystyle\sum_{k=0}^{m} {}_m\mathrm{C}_k x^k \sum_{l=0}^{n} {}_n\mathrm{C}_l x^l$ において，

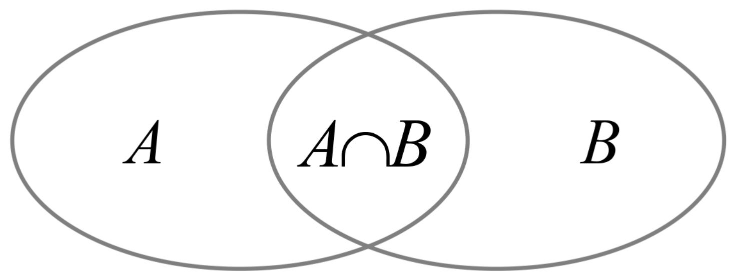
    x^r の係数を比較します.

（注）記号 $\because$ は,「なぜならば」を意味します.

包除原理 = 包含と排除の原理

- 数え上げに用いられる，集合の要素の数に関する原理です．
- S が有限集合のとき，S の要素の数を $|S|$ で表します．
- 有限集合 A, B に対して，

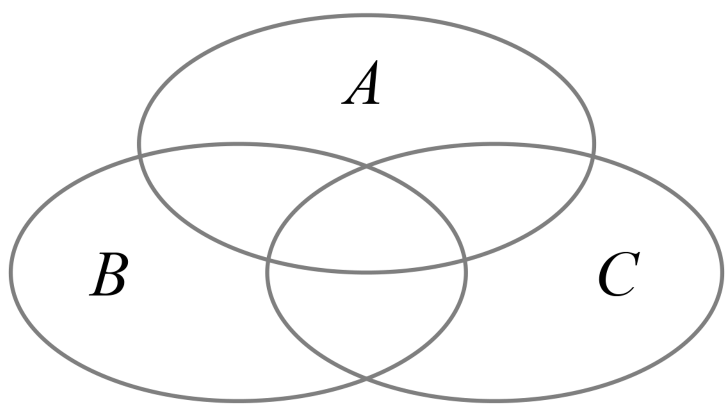
$$|A \cup B| = |A| + |B| - |A \cap B| \, .$$



包除原理 = 包含と排除の原理

- 数え上げに用いられる，集合の要素の数に関する原理です.
- S が有限集合のとき，S の要素の数を $|S|$ で表します.
- 有限集合 A, B, C に対して，

$$
\begin{aligned}
|A \cup B \cup C| = |A| + |B| + |C| \\
&- |A \cap B| - |B \cap C| - |C \cap A| \\
&+ |A \cap B \cap C| .
\end{aligned}
$$



包除原理の応用

- オイラー関数：自然数 n に対して，n と互いに素である 1 以上 n 以下の自然数の個数を $\varphi(n)$ で表します.
  - 例： $\varphi(6) = 2$ （6 と互いに素なのは，1, 5）.
    $\varphi(7) = 6$ （7 と互いに素なのは，$1, \ldots, 6$）.
    $\varphi(8) = 4$ （8 と互いに素なのは，1, 3, 5, 7）.

- $\varphi(504) = ?$
  - $504 = 2^3 \times 3^2 \times 7$ なので，504 以下で，2 の倍数でも 3 の倍数でも 7 の倍数でもない自然数の個数を求めます.
  - A を 2 の倍数，B を 3 の倍数，C を 7 の倍数の集合とすると，
    $$\phi(504) = 504 - |A \cup B \cup C|\,.$$
  - $|A| = 504/2 = 252\,, \quad |B| = 504/3 = 168\,, \quad |C| = 504/7 = 72\,.$
  - $|A \cap B| = 504/6 = 84$（6 の倍数の数），
    $|B \cap C| = 504/21 = 24\,, \quad |C \cap A| = 504/14 = 36\,,$
    $|A \cap B \cap C| = 504/42 = 12$ より，　$\phi(504) = 504 - 360 = 144\,.$

組合せは どこに現れるでしょうか

- 回帰分析や判別分析での変数選択
  - n 個の説明変数のうちの r 個で目的変数を説明するとすると，選び方は ${}_n\mathrm{C}_r$ 個あります.
- 文書要約
  - n 個の文からなる文書に対して，r 個の文を選んで要約とする．選び方は ${}_n\mathrm{C}_r$ 個あります.
- $2n$ 個の点からなるデータに 2-クラスタリングを適用するとき，解の候補は ${}_{2n}\mathrm{C}_1 + {}_{2n}\mathrm{C}_2 + \cdots + {}_{2n}\mathrm{C}_n$ 個あります.
- （注）以上の例では，すべての組合せを試してみることは，実際にはしません.
- 2 項分布の確率関数 $P(\{X = k\}) = {}_n\mathrm{C}_k p^k (1-p)^k$.

  ➪ スライド 68