数の道

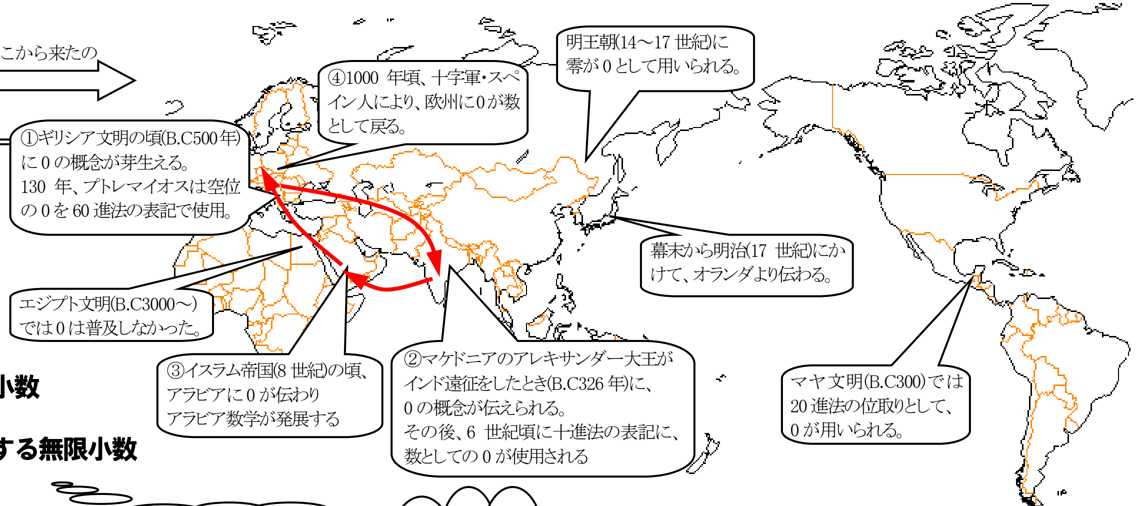
自然数(N)

原始時代より、人には指折り数え、「たくさん」略奪する和の概念は、あった。

0(零)

0 はその形から悪魔を呼び出す魔法陣の出入口に似ており、呪術的な数と見られた。
また、どんな数とも積が 0 になる不合理な数でもあり、疎んじられ広まることはなかった。欧州では、2階,1階の下は地下1階であり、0 階が抜け(0 階を地上階ということもある)、その名残がある。

0 はどこから来たの



負の自然数

やがて集落が生まれ、奪うことだけではなく与えることの必要性を人は知り負の概念が生まれる。
しかし、負数が用いられるようになったのは 6 世紀のインドであり、その後、15 世紀になってやっと欧州に広まる。

整数(Z)

国が作られ、人口が増えることで、大きな数を処理するために乗法の計算が導入される。
整数までの数は、加減乗について閉じている。

社会が形成され、交易が始まると、得た利益の分配をしなければならない。
四則演算の除法がここで生まれる。

分数

小数
- 有限小数
- 循環する無限小数

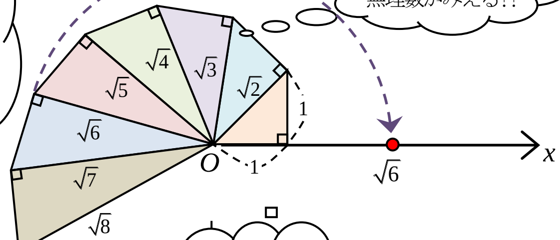
循環しない無限小数って？

有理数(Q)

無理数

実数全体は、数直線上に配置することができる。無理数は、その長さは目で見ることはできるが、正確な値は小数では表現できない。

コンパスを回せば無理数がみえる!!



1857 年、ヘンリー・リンド(スコットランド)は、エジプトのナイル湖畔のルクソールという村で、B.C1700 年頃のアーメスのパピルスを発見する。
後にリンド・パピルスといわれるこの書物は、世界最古の数学のパズル本であり、1 次方程式・不定方程式の解法が触れられており、その解は有理数の範囲で求められる。
しかし、2 次方程式の解は、有理数範囲では求めることができないものもある。

1 辺の長さが 1 である正方形の斜辺の長さ
$x^2 = 2 \Rightarrow x = \sqrt{2}$

実数(R)

ピタゴラスは B.C500 年頃のギリシアの数学者であり、ピタゴラス学派という教団を作る。
学派は「万物は数である」ことを教義とし、すべての数は細かい粒子が結びついて構成されていると考えた。「三平方の定理」は学派が発見した有名な定理であるが、その斜辺の長さは有理数で与えられないものもある。このことは学派の教義を否定することになるため、上層部は「秘密にせよ(アロゴン)」と固く口止めをし、門外秘とし、その事実を漏らしたものは暗殺されたという。しかし皮肉にも学派のシンボルである五芒星形にも黄金比という無理数は潜んでおり、やがて、この学派は市民により追放されるのである。

無理数は、2 次方程式の解として得られる。
有理数を係数とする高次方程式の解を代数的数というが、無理数の中にはこのような方程式からは求められない不思議な数もある。その代表的なものとして、円周率 π や、ネイピア数 e といったものが知られており、これらの数を超越数という。

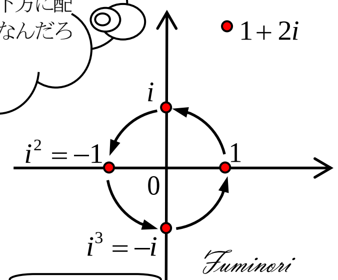
そして、神は降臨し、円周率 π を人間に与えた。

虚数

すべての実数が数直線上にあるなら、直線の上方や下方に配置した数はなんだろう?

$i^2 = -1$ ということは
$i = \pm\sqrt{-1}$　愛は負数を包む

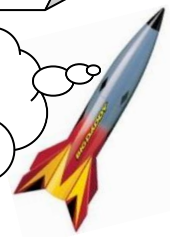
ガウス平面



複素数(C)

16 世紀、イタリアの数学者カルダノは 3 次方程式の一般解を求めようとする過程で、「平方すると負になる数」が必要であることに気づく。17 世紀になって、直交座標を考案したデカルトは、このような実在しない数を虚数とよぶ。1831 年、ガウスは実数全体が数直線上に配置できるならば、その上にこの実在しない数を配置することで、虚数の存在を示した。
　実数に虚数を合わせた複素数を用いることで、流体力学の基本概念が説明でき、人類は空そして宇宙への一歩を踏み出すのである。

複素数は、ロケットを宇宙に飛ばせるのだ。



1843 年、ハミルトン(アイルランド)は複素数を拡張した四元数を考案する。四元数は、4 次元内で視覚化することができる。
この数の概念はコンピュータの 3D グラフィックスなど IT の先端技術に利用されている。
そして、八元数が考案され、数の道はさらに延びていく。

ハミルトン数

四元数

ケーリー数

八元数

Fuminori Nakamura

完全平方式の旅

項数を増やす

$$a^2+b^2+2ab=\boxed{(a+b)^2}=a^2+2ab+b^2$$

次数を増やす

2乗して2数を掛けて2倍

３項

$$(a+b+c)^2$$
$$=a^2+b^2+c^2+2ab+2ac+2bc$$
$$=a^2+b^2+c^2+2ab+2bc+2ca$$

$a \rightarrow b \rightarrow c \rightarrow a$

数学は形式美を重んじる。
$a \to b \to c \to a \to \cdots$
のような文字の並びの式を
「輪環式」
という。

４項

$$(a+b+c+d)^2$$
$$=a^2+b^2+c^2+d^2$$
$$+2ab+2ac+2ad$$
$$+2bc+2bd$$
$$+2cd$$

2数の積は、
a を含むものを考え、次に a を含まないで b を含むものというように求めていく。

３次

$$(a+b)^3=a^3+3a^2b+3ab^2+b^3$$

４次

$$(a+b)^4=a^4+4a^3b+6a^2b^2+4ab^3+b^4$$

a は降べきの順、
b は昇べきの順
各項の係数は、
パスカルの三角形
で得られる。

1
1 1
1 2 1
1 3 3 1
1 4 6 4 1
1 5 10 10 5 1

哲学者・数学者・物理学者・神学者といったマルチ学者であったパスカルが考案したもの。
彼は、つぎの　有名な言葉を残している。
人間は考える葦である（パンセ）

Fuminori Nakamura

無理数などの計算はこの方法を用いるとスムーズな計算ができる。

$$\left(\sqrt{5}-\sqrt{3}\right)^2=(5+3)-2\sqrt{5}\times\sqrt{3}$$
$$=8-2\sqrt{15}$$

これを展開とみると、因数分解は2重根号を外すことを意味する。

$$\sqrt{p\pm2\sqrt{q}}=$$ 和 p、積 q である2数 $a,b\ (a>b)$

$$=\sqrt{a}\pm\sqrt{b}$$ （複号同順）

$$\sqrt{8-2\sqrt{15}}=(\text{和}8,\text{積}15\Rightarrow5,3)$$
$$=\sqrt{5}-\sqrt{3}$$

式の中の2文字･3文字を相互に入れ替えても元の式と変わらないような式を対称式という。対称式は基本対称式で表すことができる。
2文字の基本対称式　$a+b,\ ab$
3文字の基本対称式　$a+b+c,\ ab+bc+ca,\ abc$

$$a^2+b^2=(a+b)^2-2ab$$
$$a^2+b^2+c^2=(a+b+c)^2-2(ab+bc+ca)$$
$$a^3+b^3=(a+b)^3-3ab(a+b)$$

2次式から、完全平方式を作り出す操作を
「平方完成」
という。平方完成は、
2次関数の概形を調べるために用いられる。

$$x^2+2ax=(x+a)^2-a^2$$

$2a=p$ とおくと、

$$x^2+px=\left(x+\frac{p}{2}\right)^2-\left(\frac{p}{2}\right)^2$$

半分　引く　２乗

放物線の姿態

放物線には３つの顔がある。顔（式）は三者三様。だけど、どんなに着飾っても体型（グラフの開き）は変わらない!!

標準形

…あなたは素直。すべてがみえてしまう。

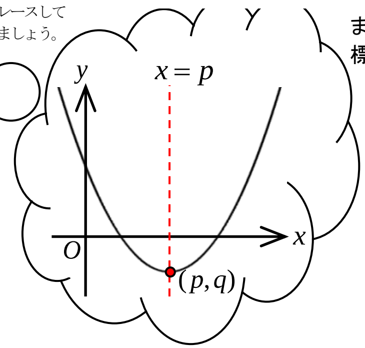
左からトレースして読み取りましょう。

$$y=a(x-p)^2+q$$

グラフの開き
$a>0$ のとき下に凸
$a<0$ のとき上に凸

q は
頂点の y 座標

$x-p=0$ が軸の方程式
このとき p は頂点の x 座標



まずは、
標準形からお付き合い

一般形

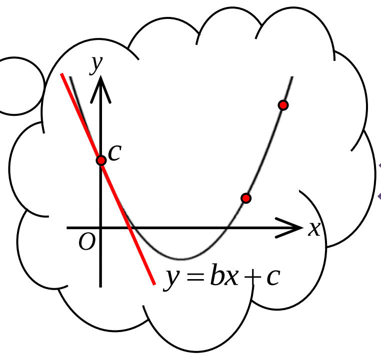
…あなたは綺麗。でも親しみにくい。

$$y=ax^2+bx+c$$

グラフの開き

y 切片

y 切片における
接線の方程式



平方完成

因数分解

切片形

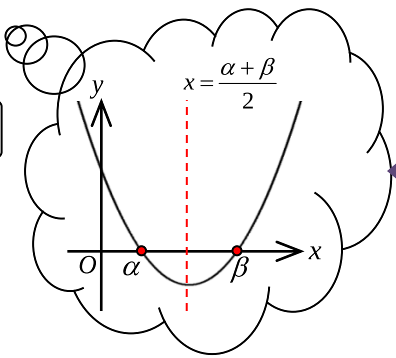
…あなたは個性的。惹かれてしまいそう。

$$y=a(x-\alpha)(x-\beta)$$

グラフの開き

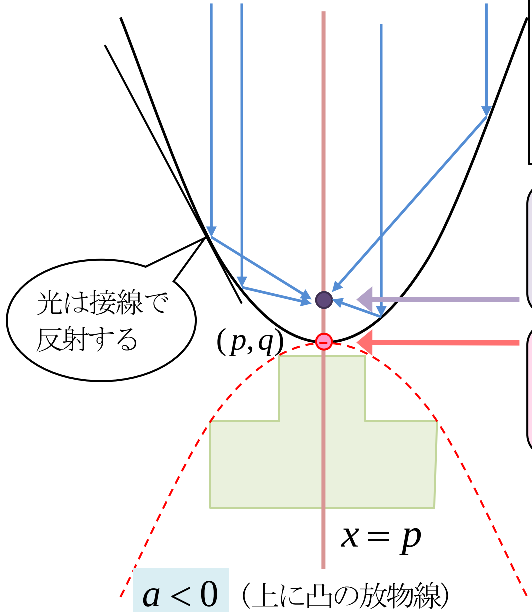
x 切片の座標は
$(\alpha,0)$，$(\beta,0)$

軸の方程式は、$x=\dfrac{\alpha+\beta}{2}$



①放物線には２つの心がある?!

$a>0$（下に凸の放物線）



$a<0$（上に凸の放物線）

パラボラアンテナは、その断面である放物線(Parabola)の軸に平行に入射した電波が焦点に集まる性質を利用したもの。
　逆に焦点からでた光は、放物線（面）に反射すると軸に平行に照射されます。この原理は、懐中電灯（車のヘッドライト）に用いられている。

焦点(Focus)
放物線の中にある
真心（愛）という心

頂点(Vertex)
放物線の下にある
下心（恋）という心

Fuminori
Nakamura

②放物線の決定

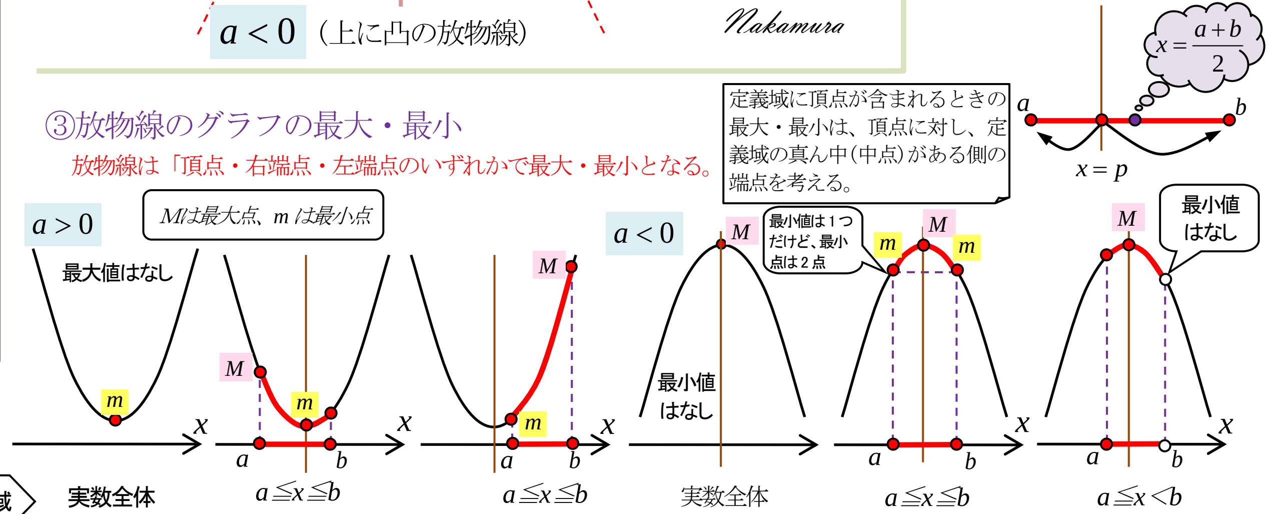
与えられた条件（性質）を満たす放物線の方程式を求める。
　一番 fit する放物線の姿態を image してみよう。
　放物線固有の性質をもっとも有するのは標準形。したがって Key-Word が
　(A) 軸・頂点 ⇒標準形
であるとき用いる。
　得体のしれない一般形は、固有の性質だけから求めることは難しい。
　(B) 2 点〜3 点 ⇒ 一般形
であるとき用いる。
　個性的な切片形は、利用も限定される。放物線が x 軸と交わっていれば、
　(C) x 軸との交点⇒切片形
であるとき用いる。

③放物線のグラフの最大・最小

放物線は「頂点・右端点・左端点のいずれかで最大・最小となる。」

定義域に頂点が含まれるときの最大・最小は、頂点に対し、定義域の真ん中（中点）がある側の端点を考える。

Mは最大点、m は最小点



2次方程式の解の配置ルール ～学名「てんじくちょう」という放物線の捕まえ方

$$ax^2+bx+c=0\ (a>0) \xrightarrow{\text{グラフ化}} \begin{cases} f(x)=ax^2+bx+c & \text{(放物線)} \\ y=0 & (x\text{ 軸}) \end{cases} \xrightarrow{\text{標準形(種類)}} f(x)=a(x-p)^2+q$$

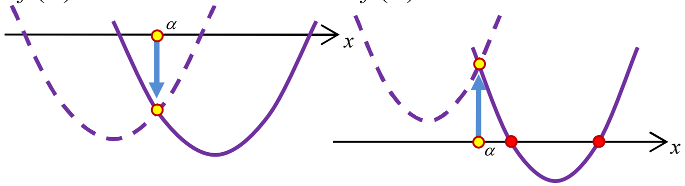
てんじくちょうの捕獲手順

チョウの2本の足が α より大きい木の枝の位置にある場合

点 — 点(Point)で、チョウを追い込め

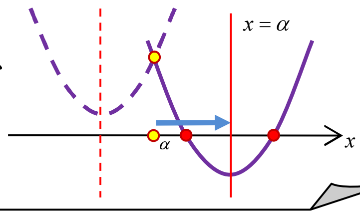
条件で与えられる点の周りしか動けないように
放物線(チョウ)を追い込む

$f(\alpha)<0$ ⇒逃げてしまう　　$f(\alpha)>0$ ⇒捕まえられるかも



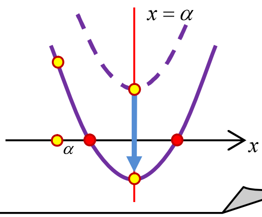
軸 — 軸(Axis)で、左右の動きを封じろ

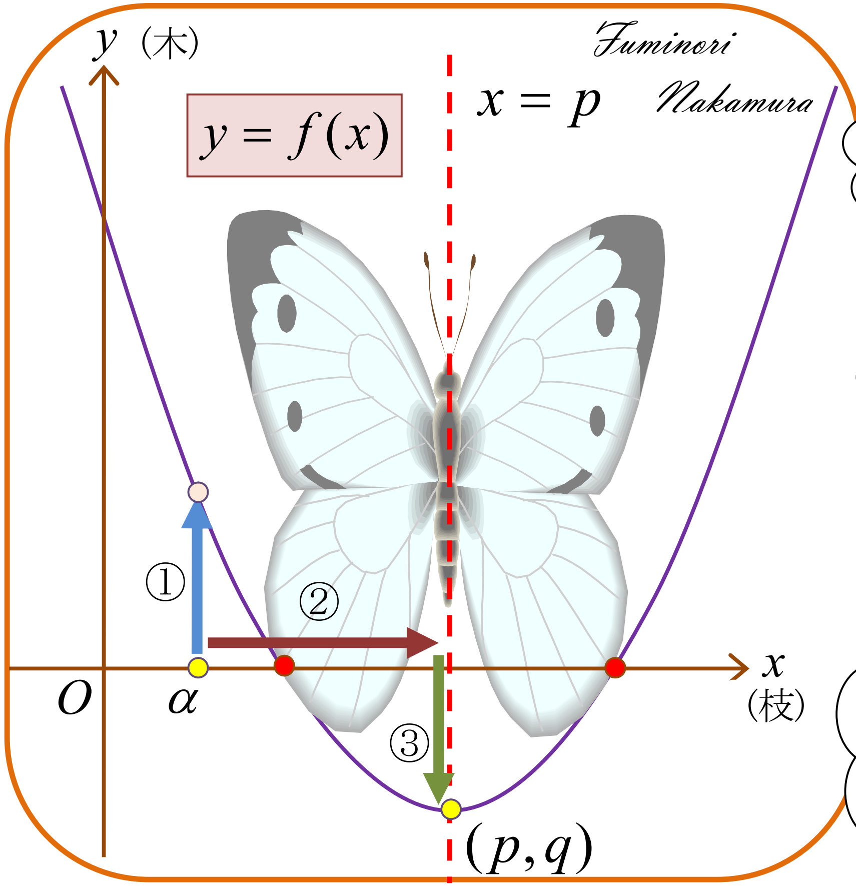
左右の移動を決定する
軸の位置の範囲を決めて、
チョウを追い詰める



頂 — 頂点(Vertex)で、上下の動きを封じろ

上下の移動を決定する
頂点の位置を上から押し下げ、
空に逃げないようにして、
チョウを捕まえる





開きが a (下に凸)、
軸 $x=p$　頂点 $(p,\ q)$
である放物線

てんじくちょう!!
と唱えて
蝶々を捕まえよう

プロは一般形から、
軸は $x=-\dfrac{b}{2a}$
x 軸と交わることは
判別式 $D\geqq 0$
として捕獲する。

てんじく蝶の標本見本

$(\alpha<\beta<\gamma)$

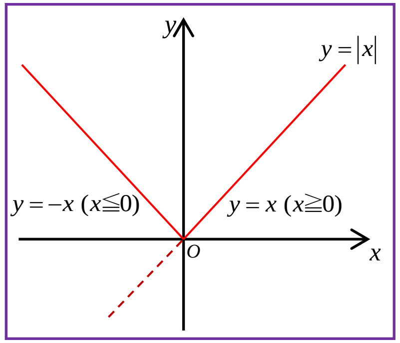
Case	解の配置	てん	じく	ちょう
A	α より大	$f(\alpha)>0$	$\alpha<p$	$q\leqq 0$
B	α より小、α より大	$f(\alpha)<0$	任意	$q\leqq f(\alpha)<0$ より必要なし
C	α と β の間	$f(\alpha)>0,\ f(\beta)>0$	$\alpha<p<\beta$	$q\leqq 0$
D	α と β 、β と γ の間	$f(\alpha)>0,\ f(\beta)<0,\ f(\gamma)>0$	必要なし	$q\leqq f(\beta)<0$ より必要なし

絶対値を含む方程式･不等式の怪

$|x|=1$ の解は、$x=\pm1$ はいいけど、$\pm x=1$ だけではダメなのだ!!

|x|が表す 3 つの性質

C x 軸上で反射する直線 $y=|x|$



A 負数を除去する *filter*

$$|x|=\begin{cases} x & (x\geqq 0)\\ -x & (x<0)\end{cases}$$

B 原点O (0) と点 $P(x)$ との距離

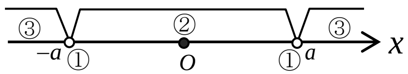
① $|x|=a \Leftrightarrow x=\pm a$

② $|x|<a \Leftrightarrow -a<x<a$

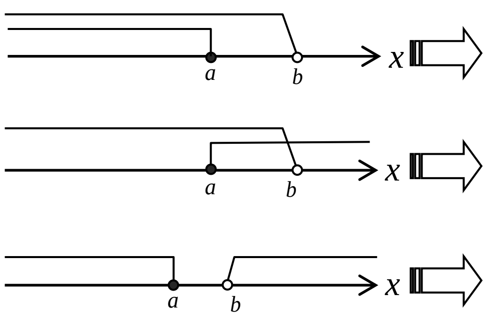
③ $|x|>a \Leftrightarrow x<-a,\ \ a<x$

カンマは、2 つの解を合わせるために用いられる。必ず書くこと

不等号の向きを優先するので、$x>a$ とは書かない。



解を合わせる(or)ことと、共通の解を求める(and)こと



解を合わせると、「$x<b$」
共通解は、「$x\leqq a$」

解を合わせると、「すべての実数」
共通解は、「$a\leqq x<b$」

解を合わせると、「$x\leqq a,\ b<x$」
共通解は、「解なし」

場合分けをするときは「解を合わせる」、
連立方程式・不等式の式を解くときは「共通の解を求める」

★☆★☆★☆☆★★

(1) $|2x-1|=3$

原点と点 $P(2x-1)$ との距離が 3 に等しい

$\pm(2x-1)=3$ だけでは不十分

B

$2x-1=\pm3$
より、 $x=-1, 2$

$2x-1<\pm3$ は日本語になっていない。
$\pm(2x-1)<3$ は、場合分けをしていない。

(2) $|2x-1|<3$

原点と点 $P(2x-1)$ との距離が 3 より小さい

B

$-3<2x-1<3$
より、 $-1<x<2$

$|1-2x|=|2x-1|$ であることを用いて x の係数を正にしよう。

(3) $|1-2x|>3$

原点と点 $P(2x-1)$ との距離が 3 より大きい

B

$2x-1<-3,\ \ 3<2x-1$
より、 $x<-1,\ 2<x$

(4) $|2x-4|<x+1$

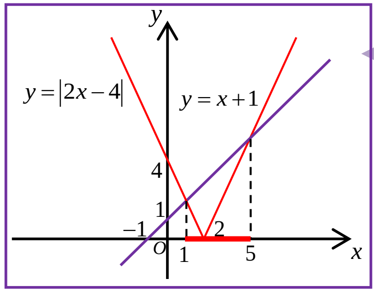
原点からの距離として式は読めない

$$|2x-4|=\begin{cases} 2x-4 & (x>2)\\ -(2x-4) & (x\leqq 2)\end{cases}$$

A 場合分けすると

①$x\leqq2$ のとき、
$-(2x-4)<x+1$ より、 $x>1$
$x\leqq2$ との<u>共通範囲を求めて</u>、 $1<x\leqq2$
②$x>2$ のとき、
$(2x-4)<x+1$ より、 $x<5$
$x>2$ との<u>共通範囲を求めて</u>、 $2<x<5$
①と②の<u>解を合わせて</u>、 $1<x<5$

C グラフにすると

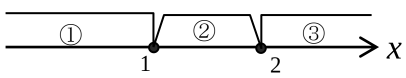


(5) $|x-1|+|x-2|<x+1$

原点からの距離として式は読めるはずもない。

$$|x-1|=\begin{cases} x-1 & (x>1)\\ -(x-1) & (x\leqq 1)\end{cases},\quad |x-2|=\begin{cases} x-2 & (x>2)\\ -(x-2) & (x\leqq 2)\end{cases}$$

A $x-1=0,\ x-2=0$ である、
$x=1,\ 2$ の前後で場合分け。



$|x-1|$ は、数直線上で $x=1$ の左にある①は、(−)をつけてはずす。
$|x-2|$ は、数直線上で $x=2$ の左にある①②は、(−)をつけてはずす。

①$x\leqq1$ のとき
$-(x-1)-(x-2)>x+1$ より、 $x>\dfrac{2}{3}$。 $x\leqq1$ との<u>共通範囲を求めて</u>、 $\dfrac{2}{3}<x\leqq1$
②$1<x<2$ のとき
$(x-1)-(x-2)>x+1$ より、 $x<0$。 $1<x<2$ との<u>共通範囲を求めて</u>、解はなし。
③$2\leqq x$ のとき
$(x-1)+(x-2)>x+1$ より、 $x>4$ 　 $x\geqq2$ との<u>共通範囲を求めて</u>、 $x>4$

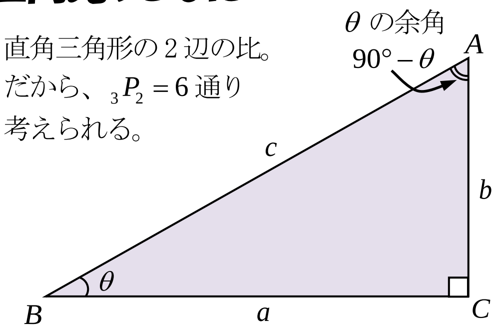
①②③の<u>解を合わせて</u>、 $\dfrac{2}{3}<x\leqq1,\ \ 4<x$

Fuminori Nakamura

三角比の坂 ～ 三角比の歴史坂を駆け抜けよう!

三角比ってなに？

直角三角形の２辺の比。
だから、$_3P_2=6$ 通り
考えられる。



逆数が等しい三角比は１つと考え、この３つの比を三角比とする。

$\frac{a}{c}=(\text{余弦})=\cos\theta=\sin(90^\circ-\theta)$

$\frac{b}{c}=(\text{正弦})=\sin\theta$

$\frac{b}{a}=(\text{正接})=\tan\theta$

$\frac{c}{a}=(\text{正割})=\sec\theta$

$\frac{a}{b}=(\text{余接})=\cot\theta=\tan(90^\circ-\theta)$

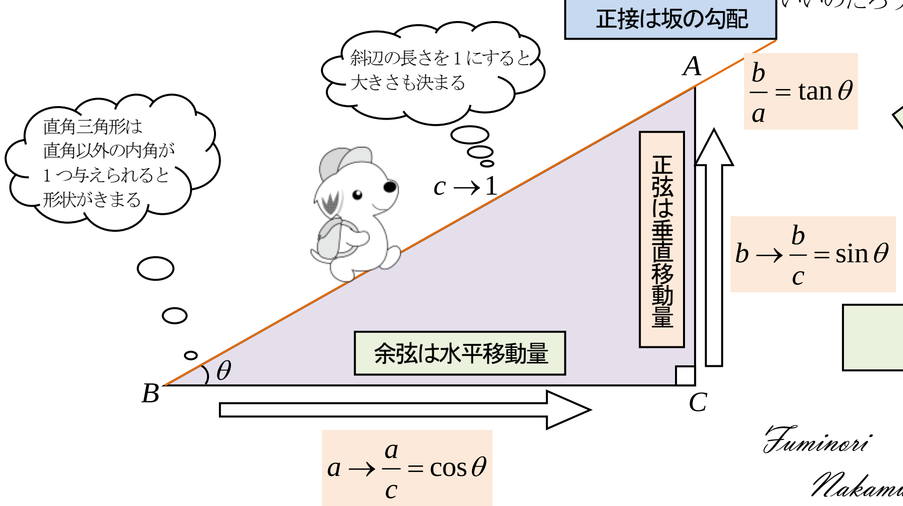
$\frac{c}{b}=(\text{正割})=\csc\theta=\sec(90^\circ-\theta)$

6つの比の値は、正 or 余 {弦 接 割} と表す。

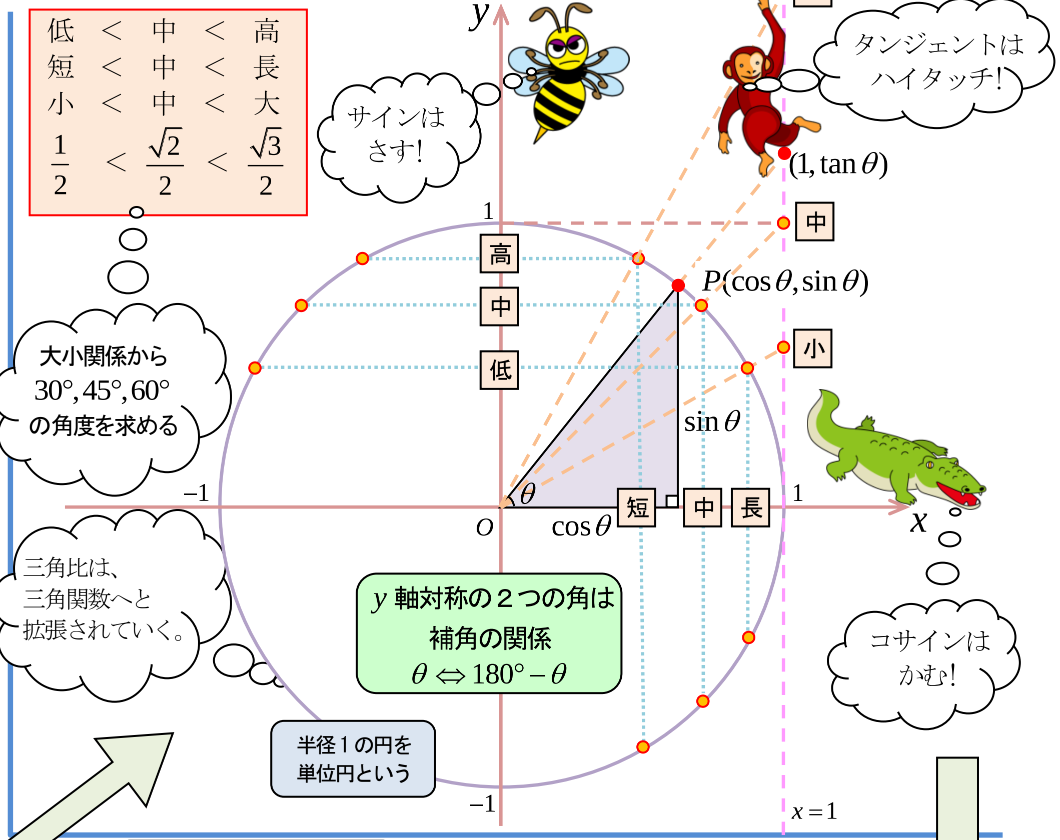
「余」がつく三角比は、「正」の角 θ に対して余角 $(90^\circ-\theta)$ の正弦・正接・正割を示す。記号では co (with) を接頭語としてつける。

３つの三角比
サイン、コサイン、タンジェント
の意味を考えてみよう！

坂を登ると三角比が見える!!!



90° 以上の角の三角比はどうやって求めたらいいのだろう？



三角比の相互関係

…１つの三角比から残りの２つを求めてみよう

三角比の基本性質
(三平方の定理のこと)

- $\cos^2\theta+\sin^2\theta=1$
- $\tan\theta=\frac{\sin\theta}{\cos\theta}$
- $1+\tan^2\theta=\frac{1}{\cos^2\theta}$

両辺を $\cos^2\theta$ で割る

基本性質の三態

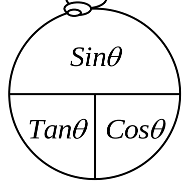
① $(\cos\theta+\sin\theta)^2=1+2\cos\theta\sin\theta$
② $\sin^2\theta=1-\cos^2\theta$
③ $\cos^2\theta=1-\sin^2\theta$

$\sin\theta=\cos\theta\tan\theta$

$\tan\theta$ が与えられたとき、残りの三角比を求める手順。

$\cos^2\theta=\frac{1}{1+\tan^2\theta}$

坂で転んでスッテンコロリン!



Fuminori Nakamura

三角形の辺と角の相性 ～三角形の上で、三辺対三角のフィーリングカップルを探してみよう！

Fuminori Nakamura

三角形の存在条件

① $a>0,\ b>0,\ c>0$　（辺の長さは正だよ）

② 三角形の１辺の長さは他の２辺の長さの和より小
（犬の習性という法則　…　菊池寛曰く）

$a<b+c$
$b<c+a$
$c<a+b$
⇒ 最大辺の長さが c のとき $c<a+b$

辺と角の関係

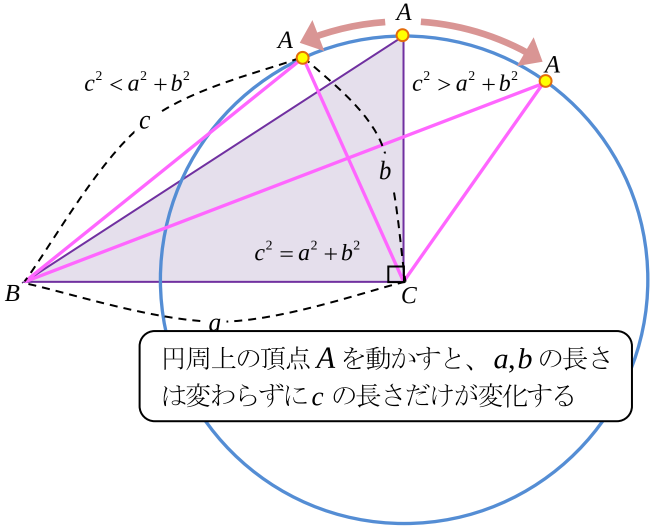
$A<B<C \iff a<b<c$ ⇒ 最大(最小)角　対辺⇓⇑対角　最大(最小)辺

三角形の形状

$$\cos C=\frac{a^2+b^2-c^2}{2ab}$$

最大辺の長さを c とすると、

① 鋭角三角形 ⇔ $c^2<a^2+b^2$　$\Leftarrow \cos C>0$
② 直角三角形 ⇔ $c^2=a^2+b^2$　$\Leftarrow \cos C=0$
③ 鈍角三角形 ⇔ $c^2>a^2+b^2$　$\Leftarrow \cos C>0$

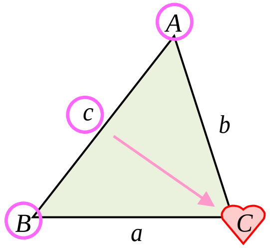


円周上の頂点 A を動かすと、a,b の長さは変わらずに c の長さだけが変化する

三角形の決定条件と解法

解法のPoint
- 向かい合う辺と角を正弦・余弦定理を用いて１組求める。⇒　♥カップル成立
- １組のカップルを仲人にして他の辺と角を求める。
  最初は正弦定理でアプローチ。
  ダメな場合は、成立した１組のカップルの関係から第２余弦定理でアプローチ。

① １辺と２つの角が与えられている　$(A,\ B,\ c)\Rightarrow(a,\ b,\ C)$



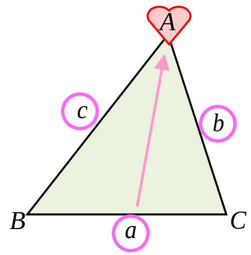
内角の和：$C=180^\circ-A-B$

C ♥ c ⇒ 正弦定理：$a=\dfrac{c\sin A}{\sin C}$ ⇒ 第１余弦定理：$b=c\cos A+a\cos C$

Cの三角比の値が求められる。

Cの三角比の値が求められない。 ⇒ 正弦定理：$a=2R\sin A,\ b=2R\sin B$ ⇒ 第１余弦定理：$c=a\cos B+b\cos A$ より R を求める

② ３辺の長さが与えられている　$(a,\ b,\ c)\Rightarrow(A,\ B,\ C)$



第３余弦定理：$\cos A=\dfrac{b^2+c^2-a^2}{2bc}$

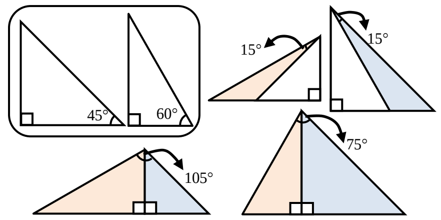
辺の値がキレイな対角を求める。

A ♥ a ⇒ 第３余弦定理：$\cos B=\dfrac{c^2+a^2-b^2}{2ca}$ ⇒ 内角の和：$C=180^\circ-A-B$

次に辺の値がキレイな対角 ⇒ 正弦定理：$\sin B=\dfrac{b\sin A}{a}$ ⇒ 内角の和：$C=180^\circ-A-B$

対辺の長さの大小関係から角Bを決める。

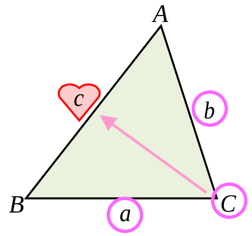
２つの直角定規を組み合わせて三角形の形が見えることがある。



だから、15°の三角比の値を覚えておくと三角比が解ける。

$$\sin 15^\circ=\frac{\sqrt{6}-\sqrt{2}}{4}\qquad \cos 15^\circ=\frac{\sqrt{6}+\sqrt{2}}{4}$$

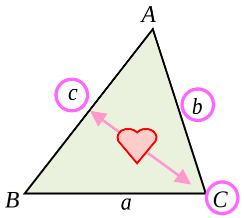
③ ２辺とその間の角が与えられている　$(a,\ b,\ C)\Rightarrow(A,\ B,\ c)$



第２余弦定理：$c^2=a^2+b^2-ab\cos C$

C ♥ c ⇒ 正弦定理：$a=\dfrac{c\sin A}{\sin C}$ ⇒ 内角の和：$C=180^\circ-A-B$

④ ２辺とどちらかの辺の対角が与えられている　$(b,c,\ C)\Rightarrow(a,\ A,\ B)$　← 三角形は必ずしも決定はしていない



三角比の値から B が求められる
⇓
b と c の大小関係から B の値は２つ得られる場合もある。

正弦定理：$\sin B=\dfrac{b\sin C}{c}$ ⇒ 内角の和：$A=180^\circ-B-C$ ⇒ 正弦定理：$a=\dfrac{c\sin A}{\sin C}$

a の２次方程式を解く
⇓
a の値は２つ得られる場合もある。

第２余弦定理：$c^2=a^2+b^2-2ab\cos C$ ⇒ 正弦定理：$\sin A=\dfrac{a\sin C}{c}$ ⇒ 内角の和：$B=180^\circ-C-A$

データの散らばりの数値化

変量 X のデータ

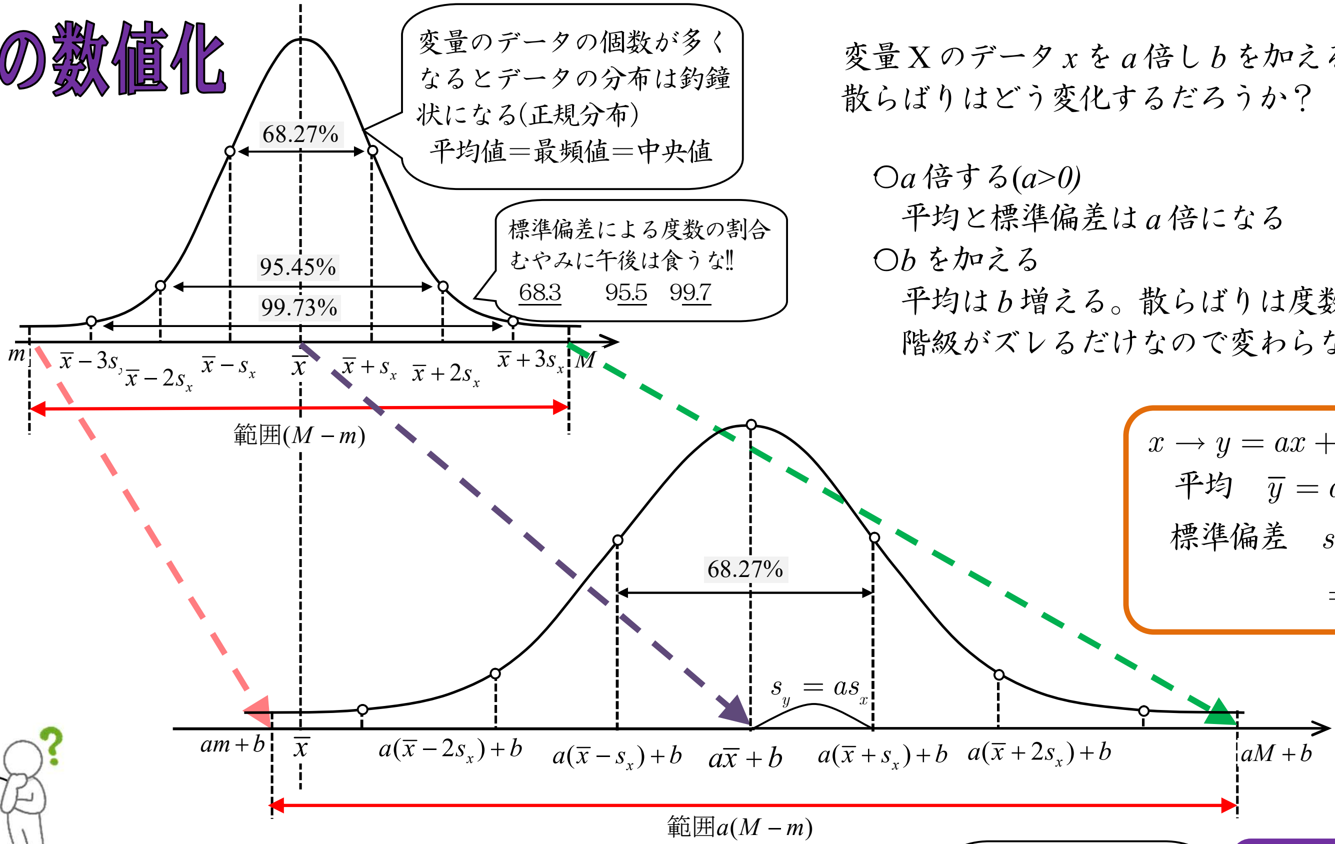
$x : x_1, x_2, x_3, \cdots\cdots, x_n$

の散らばりの数値化ステップ

変量 X のデータ x を a 倍し b を加えると，
散らばりはどう変化するだろうか？

Fuminori Nakamura

○a 倍する$(a>0)$
　平均と標準偏差は a 倍になる
○b を加える
　平均は b 増える。散らばりは度数分布の
　階級がズレるだけなので変わらない。



$x \to y = ax+b$ とすると
　平均　$\bar{y} = a\bar{x}+b$
　標準偏差　$s_y = |a|s_x$
　　$\Rightarrow s_y^2 = a^2 s_x^2$

<Step1>
データの平均 $\bar{x}$ を求める

$$\bar{x} = \frac{x_1+x_2+x_3+\cdots+x_n}{n}$$

<Step2>
平均からの散らばり(偏差)

$x_1-\bar{x},\ x_2-\bar{x},\ x_3-\bar{x},\ \cdots,\ x_n-\bar{x}$

偏差の和は０より，
和で散らばりは表せない

<Step3>
散らばりを偏差の大きさの和で数値化

○絶対値で散らばりを表す

$$|x_1-\bar{x}|+|x_2-\bar{x}|+\cdots+|x_n-\bar{x}|$$

○２乗して散らばりを表す

$$(x_1-\bar{x})^2+(x_2-\bar{x})^2+\cdots+(x_n-\bar{x})^2$$

２つの変量のデータの個数が違うと，
散らばりの比較はできない。

<Step4>
データの個数で割り，散らばりを揃える(平均化)

○平均偏差(*Mean Deviation*)

$$MD = \frac{|x_1-\bar{x}|+|x_2-\bar{x}|+\cdots+|x_n-\bar{x}|}{n}$$

○分散(*Variance*)

$$s_x^2 = \frac{(x_1-\bar{x})^2+(x_2-\bar{x})^2+\cdots+(x_n-\bar{x})^2}{n}$$

分散の単位はデータの単位の２乗になってしまう
例えば，長さ(cm)の分散は面積(cm²)

偏差を計算しないで
標準偏差(分散)は
求められる

○標準偏差の簡便式

$$s_x = \sqrt{\overline{x^2}-(\bar{x})^2}$$

<Step5>
データと分散の散らばりの単位を揃える
○標準偏差(*Standard Deviation*)

$$s_x = \sqrt{\frac{(x_1-\bar{x})^2+(x_2-\bar{x})^2+\cdots+(x_n-\bar{x})^2}{n}}$$

目盛の変換
　$s \Rightarrow 1 \Rightarrow 10$
平均の変換
　$\bar{x} \Rightarrow x-\bar{x}(0) \Rightarrow 50$

○偏差値(*Standard Score*)

$$ss = \frac{x-\bar{x}}{s_x}\times 10+50$$

データの散らばりの視覚化

Fuminori Nakamura

まずは，おおざっぱに資料の特徴を視る

代表値

資料全体の中心位置を表す指標

中央値(Median)
　変量を小さい順に並べた中央の位置の値
（偶数個のときは，中央の２つの値の平均）
最頻値(Mode)
　変量で度数が最大である値
（度数分布表では，最大度数の階級の階級値）
平均値(Mean)
　変量の値の総和を変量の個数で割った値

３つのＭ
さらに，
最大値(Max)
最小値(Min)
を含めると，
５つのＭ

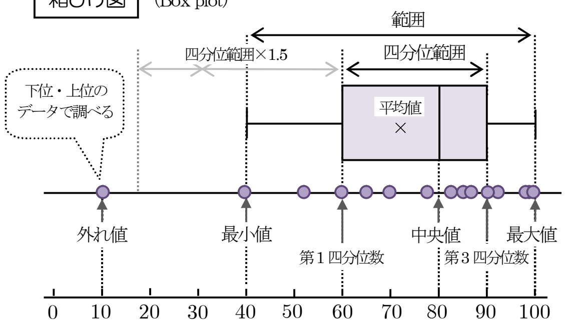
１次元データの整理・視覚化

変量・データ

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	平均
変量	10	40	52	60	64	70	78	82	85	86	90	92	99	100	72.0
区分	下位データ							上位データ							

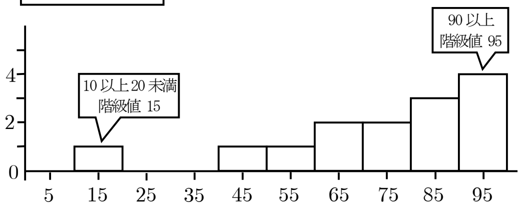
箱ひげ図

(Box plot)



ヒストグラム

(Histogram)



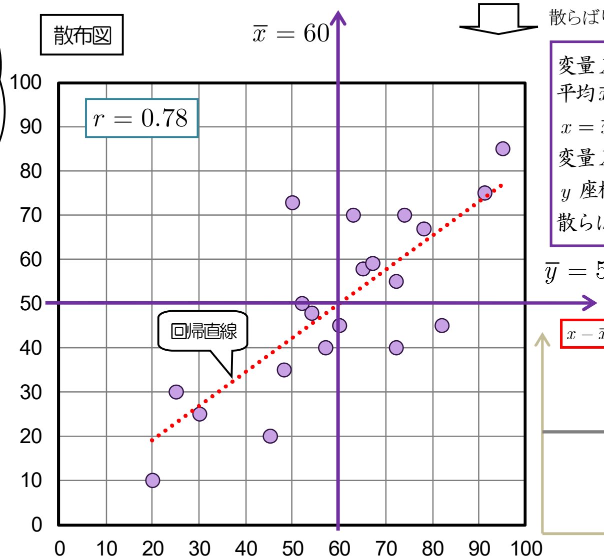
ヒストグラムと棒グラフは異なる!!!
ヒストグラムは度数分布表の階級(横)と度数(縦)の関係を視覚化した図。データは，小さい順に並んでいる連続量(棒の間隔はない)
棒グラフは，個別の項目の個数を視覚化　例えば科目別の得点の比較など。棒の間に隙間があり，小さい順に並んでなくてもいい

２次元データの整理・視覚化

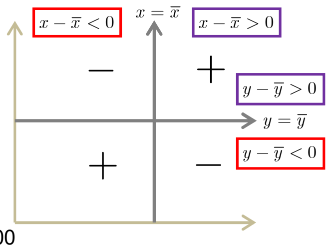
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	平均
数学X	20	25	30	45	48	50	52	54	57	60	63	65	67	72	72	74	78	82	91	95	60.0
理科Y	10	30	25	20	35	73	50	48	40	45	70	58	59	55	40	70	67	45	75	85	50.0

散らばりの視覚化

散布図



変量 X と変量 Y のそれぞれの平均 $\overline{x}, \overline{y}$ を計算し，新座標軸 $x = \overline{x}$，$y = \overline{y}$ を作成する。
変量 X, Y の偏差の値を x 座標，y 座標として点をプロットして散らばりを調べる。



散らばりの数値化

変量 X と変量 Y の偏差積
　第１，第３象限の点では正の値
　第２，第４象限は点では負の値

正・負の傾向を調べる

偏差積の和で散らばりを数値化

$$\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})$$

平均をとってデータの個数を調整する

共分散 (Covariance)

$$s_{xy} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})$$

データの単位を揃える

変量 X, Y の標準偏差で割り散らばりの値を揃える

$$r = \frac{s_{xy}}{s_x s_y} \quad (-1 \leqq r \leqq 1)$$

相関係数(Correlation coefficient)

共分散のスリム化

$$s_{xy} = \frac{1}{n}\sum(x_i - \overline{x})(y_i - \overline{y})$$
$$= \frac{1}{n}\sum\left(x_i y_i - x_i\overline{y} - \overline{x}y_i + \overline{x}\overline{y}\right)$$
$$= \overline{xy} - \overline{x} \cdot \overline{y}$$

相関係数の代数的な意味

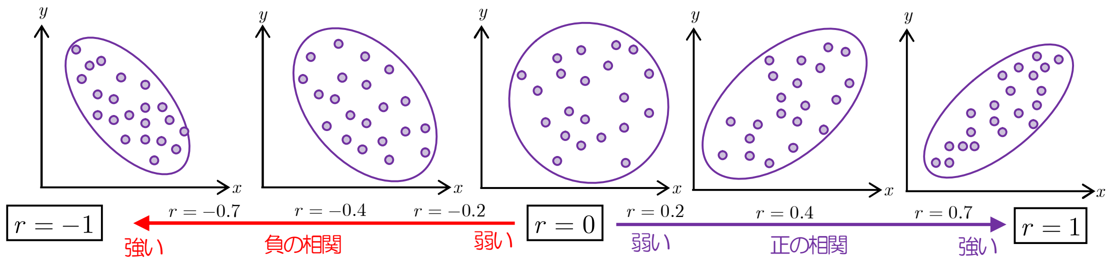
$$\vec{p} = (x_1 - \overline{x},\ x_2 - \overline{x},\ \cdots,\ x_n - \overline{x})$$
$$\vec{q} = (y_1 - \overline{y},\ y_2 - \overline{y},\ \cdots,\ y_n - \overline{y})$$

とすると，

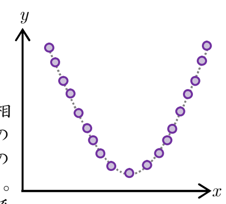
$$\vec{p} \cdot \vec{q} = \left|\vec{p}\right|\left|\vec{q}\right|\cos\theta \text{ より，}$$

$$\cos\theta = \frac{\vec{p} \cdot \vec{q}}{\left|\vec{p}\right|\left|\vec{q}\right|} = \frac{s_{xy}}{s_x \cdot s_y} = r$$

$-1 \leqq \cos\theta \leqq 1$ より，$-1 \leqq r \leqq 1$
相関係数は方向余弦になる



右の散布図の相関係数を求めるとほぼ０より相関性はないが，散布図のデータは放物線を描くので，何らかの関係はある。相関係数は直線的な関係の強さしか表せない。



回帰直線

(Regression line)

変量 X のデータを新たに加えると，対応する変量 Y のデータがおおよそどのような値になるか，推測することができる。
相関性が強く X と Y の間に比例関係がある場合，推測できる点は直線上に近似することができる。このような直線を回帰直線という。

回帰直線の傾き(回帰係数) a は，次の式で計算できる。

$$a = \frac{s_{xy}}{s_x^2} = \frac{s_{xy}}{s_x \cdot s_y} \times \frac{s_y}{s_x} = r \cdot \frac{s_y}{s_x}$$

点 $(\overline{x}, \overline{y})$ を通る直線を考えると，回帰直線は，

$$y = r \cdot \frac{s_y}{x_x}(x - \overline{x}) + \overline{y}$$

英国の生物学者ゴルトンは，親の特性は子どもに2/3しか遺伝されず，残りは種族の特性の平均値に回帰することを実験で示した。ここから回帰という用語が生まれた。

記述統計から推測統計へ

【記述統計】…全体から部分を分析
　データを整理し，図表やグラフ(箱ひげ図・ヒストグラム・散布図等)を用いてデータの特徴や傾向を調べる統計法。
　英国を中心に 19 世紀から 20 世紀にかけて発達する。ゴルトンは相関の概念に到達し，ピアソンは相関係数を定義した。また，ナイチンゲールは，統計学の知識より作成した資料をもとに，クリミア戦争での戦死者数を激減させ統計学の母といわれるようになった。

【推測統計】…部分から全体を推論
　国勢調査のように，対象となるすべての変量を調査することは難しい。そこで，平等に抽出された一部(標本)のデータの特性(平均や分散)を調べ，推定・検定といった方法で全体(母集団)のデータの状況を推論する統計が 1920 年頃から研究が始まった。確率論の発達がその研究に大きな影響を与えた。