

## 境界・筆界実務における誤差の基本

実務において「長さの誤差はどのぐらいですか。」「長さの精度はどのぐらいですか。」という言い方をします。一見して同じ内容に聞こえますが実は違います, このことをしっかり理解しておかないと誤解の元になりますので注意が必要です。

境界(筆界)復元に於いては誤差論, 確率論, 最小二乗法のこの三つの考え方は密接に関係しており, 別々に説明出来るものではありません, これからの説明ではその内容によって重複して説明される部分があります。

### 平均と標準偏差(バラツキ)の関係

「長さの誤差はどのぐらいですか。」とは (1)平均で (2)中間値で (3)大きくても (4)精度(標準偏差)で のどれを指すのでしょうか。

(1) 平均でどのぐらいかということであれば, 10個のデータがある場合でいえば, 測定データ(通常は自分が測った数値)と図面データ(図面等を書いてある数値)の差を誤差というのです, 誤差には+もあれば-もありますので

誤差を  $x_1 x_2 x_3 \cdots x_{10}$  とすれば

No.	$x_i$
1	5
2	4
3	-10
4	-1
5	-5
6	1
7	12
8	-9
9	3
10	-8
(1)平均	-0.8
(2)中間値	6
(3)大きい値	12
(4)精度(標準偏差)	7.1

(1) 誤差の平均値  $\bar{x} = \frac{x_1 + x_2 + x_3 + \cdots + x_{10}}{10}$  で計算されることとなります。表では -0.8 となります, 通常はプラスとマイナスは数, 数量とも同程度に発生しますので誤差の平均値  $\bar{x}$  は 0 に近づいてしまいます。

(2) 中間値は 0 とプラスの最大値もしくはマイナスの最小値の中間値で表しますので表では  $\bar{x} = \frac{12}{2} = 6$  を言います。

(3) 大きい値 は表では 12 です。

(1)(2)(3)で実際に計算していないで経験的という場合があります, 「昭和40年代の平板測

量に於ける測定距離の誤差はどのくらいですか。」この場合は経験的な数値を感覚的に言っていることがほとんどですが数値は(2)中間値か(3)大きい値を言っています。

人によっては「測量のプロが測っていれば 3cm ぐらいで素人でも 5cm ぐらいです。」あるいは「30m 程度なら 3cm ぐらいで 50m 程度なら 5cm ぐらいです。」という言い方をします。受け取る側はどの数値と解釈しているかは解りません。

(3) 標準偏差は測定データを  $x_1 x_2 x_3 \cdots x_{10}$  とすれば

$$\text{平均値 } \bar{x} = \frac{x_1 + x_2 + x_3 + \cdots + x_{10}}{10}$$

データ数  $n$

$$\text{標準偏差 } \sigma = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2 + \cdots + (x_{10} - \bar{x})^2}{(n-1)}} \text{ で計算され、表では 7.1 です。}$$

「器械の距離測定精度はどのくらいですか。」とは標準偏差のことです。標準偏差は誤差のバラツキ具合を示す指標です。測量器械の精度は $\pm 3\text{mm}$ (器械のカタログには標準偏差の値が書かれています。)とあれば平均では  $3 \times 0.674 = 2.0\text{mm}$  で、平均値で考える方は「2mm ぐらいの誤差があるのか。」と理解しなければなりません。0.674 は確率 50% の標準偏差の値です。

「誤差は幾らか。」という場合と「精度は幾らか。」という場合では数値のもつ意味が違います、このこと知って理解できるようになることが大事です。

### 標準偏差と平均二乗誤差と平均値の関係

平均二乗誤差は国土調査法施行令別表4にある指標で、簡単に言えば抜き取り検査結果の合格、不合格を判断する指標です、統計的な指標ではありませんので平均二乗誤差から統計的な判断をする場合は平均二乗誤差を一旦標準偏差に変換する必要があります。平均二乗誤差は一組の座標値(X値, Y値)の二変量から地籍座標値と検査値から乖離(点間距離)を求めた数値ですから比較するには二変量標準偏差に変換して比較する必要があります。

標準偏差と平均二乗誤差の違いの説明は他の項でしていますのでそちらを参考にして頂くとして二変量標準偏差 $\times \sqrt{2} =$  平均二乗誤差 の関係にあります。

平均二乗誤差が 3cm の場合の標準偏差は  $\frac{3}{\sqrt{2}} = 2.1\text{cm}$  です。

この関係から次ことがいえます

平均二乗誤差の 3 倍は公差になっていますので、公差の  $3 \times \sqrt{2} = 4.24$  分の 1 が標準偏差となります。これは確率 10000 分の 9999 に相当します。

このことから、国土調査法施行令別表4 にある距離の公差、面積の公差も 10000 分の 9999 と考えれば良いこととなります、距離の公差、面積の公差は一変量ですから 3.891 倍

標準偏差となります。

仮に距離誤差の公差が200mmとしたときに  $\frac{200}{3.891} = 51.40$  が最低限必要な精度(標準偏差)になります。通常遭遇する確率とは95%(有意水準5%)を言いますので、確率95%は1.96倍標準偏差ですから、 $51.40 \times 1.96 = 100.74$ 、通常遭遇する値は100.74を超えることはあり得ないと考えるのです。これは公差の約2分の1と考えれば解りやすいです。

大雑把に公差とか許容誤差の約2分の1を超える値に遭遇したら“これはチョット怪しいぞ”と覚えることが大事です。

筆界特定書、鑑定書とか判決に“公差内であるから境界(筆界)を特定する数値として採用できる”といった表現が使われますがこの使い方は適切ではありません。公差とは最悪でもこの範囲以内なら許される巾を決めているもので、最悪ですから通常遭遇する数値ではありません。

公差をこのような判断に使う場合はおおよそ公差を4等分して、4分の1以内は通常遭遇する範囲、4分の2以内は頻繁ではないが遭遇する範囲、4分の3以内なら遭遇することが希な範囲、4分の4以下なら非常に希な範囲で通常ではあり得ない範囲という認識が必要です。それぞれの範囲のよって必要な処置を取ったうえでどのような使い方をするかを判断しなければなりません。

それでも、境界測量をしている士業者は少なくとも「平均と標準偏差(バラツキ)の関係」は知っていないとこまります。

### 一般的な誤差の解説

誤差は一般的に定誤差(系統的誤差とも言う)、過失誤差、偶然誤差の3つがあり、観測値から真値を差し引いた値が誤差であるが真値を知ることは出来ないので真値に近い値、最も確からしい値、最確値を求め真値に代用すると言った説明がされていると思います。

#### 定誤差(系統的誤差)

機器誤差：測定機器の校正不良、補正不良、或いは機器が正しく設置されていないことによるずれなどがあります。また測量機器によって測角、測距の正確さが違いますし個々のもっている機器ごとにばらつきがあるものです。

理論誤差：観測値というのは直接測定して得られるものだけとは限りません。直接は測れないけれども、すでに測定した別の量の値を組み合わせることで理論的に得られる値もあります。その時に使った理論、計算式が違っていたならばこの誤差は理論を修正することで減らすことができるはずです。

個人誤差： 目標を視準するさいに右寄りに合わせる人、左寄りに合わせる人、もしかしたら乱視気味の人と色々います、これらは大抵の場合無意識です。その量が少ない場合の個人誤差はどちらかと言えば偶然誤差に分類されるべきかもしれません。

#### 過失誤差

観測者の勘違いや機器の操作ミスなどによって入ってくる誤差。目盛りの読み間違い、記録間違い、観測に必要な条件が正しく整っているかを事前に確認するのを忘れたなどで起きます。今では観測値は電子的に記憶されますので目盛りの読み間違い、記録間違いはないでしょうがそれでも過失が起り得ることを無視してはいけません。

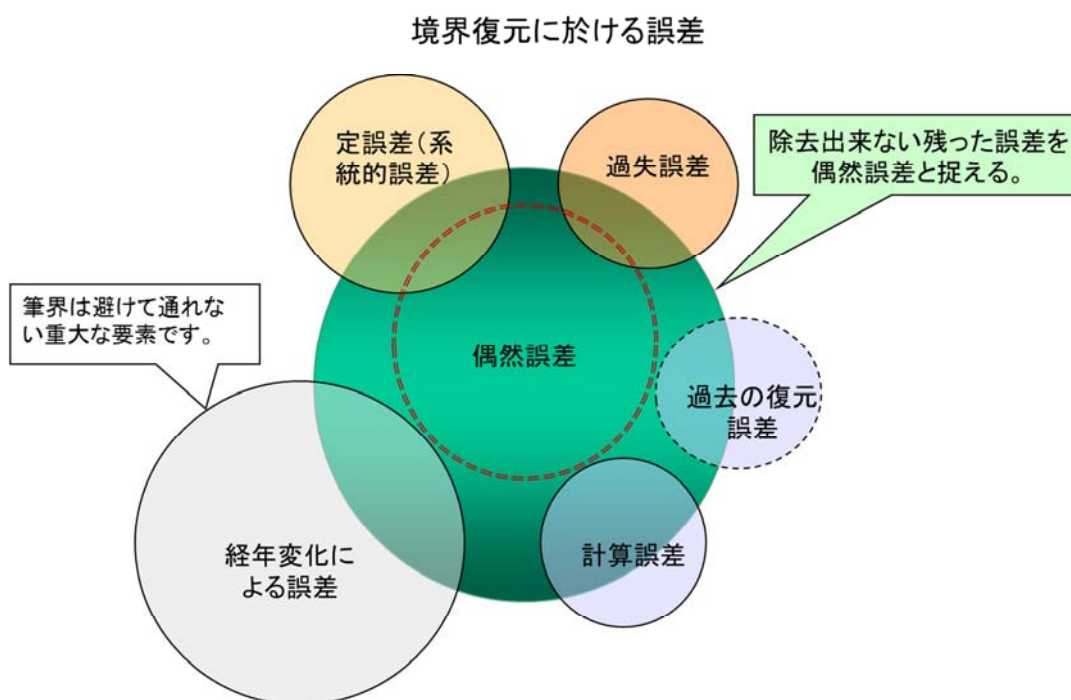
#### 偶然誤差

測定誤差： 同じ方法で観測しているように見えて、観測のたびに条件が微妙に変わったりすることは良くあります。観測者する人間の五感も一定ではありません。観測方法の中に読み取り値の変動を許す要因が含まれている場合があります。

統計誤差： 測量機器の性能 (TSの性能と誤差の発生量を参照)、精度は機器によって5秒精度とか10秒精度と異なります、この場合制御できない要因の組み合わせによって測定値にバラつきが出ます。さらに器械の状態も、使用環境も常に変化しています。

#### 境界(筆界)復元に於ける誤差とは

誤差のイメージは下図のようになります、これまで説明してきた定誤差、過失誤差以外にも境界(筆界)復元では復元計算に大きな影響を及ぼす誤差があります。



それは過去の境界復元誤差, 測量の計算誤差, 境界標の経年変化による誤差, 経年誤差があります。他にも色々ありますが大きく見ればこの三つでしょう, 上図にある一点鎖線の偶然誤差が本来の大きさとすれば時間の経過とともに様々な誤差を吸収して外側の実線円(緑の円)の大きさになります, このてんが境界(筆界)復元を難しくしているのです。

#### 過去の復元誤差

境界(筆界)を測量した図面の数値的根拠を示さず, 単に不正確という判断の下に現況占有物を基準に関係者の合意を得ることを条件に境界(筆界)復元をするということが長い間行われてきました, このことによる復元誤差が存在します。

又, 境界(筆界)復元で使われるデータはX・Y座標値でこの座標値は図面成果と実測成果を比較して計算をすすめるものですが図面成果に座標値が記載されてい無い場合はそれを下に画地調整によって得られた値を使うことが求められます, がそれをせずに境界(筆界)復元が行われてきました, 今でもされています, このことによる復元誤差が存在します。

#### 計算誤差

両成果はともに距離と角度という混合測量によって得られた観測値を下に計算されたものです, 観測の時, 測量計算の段階で定誤差(系統的誤差), 過失誤差は取り除かれています。それでも図面成果と実測成果から求めた誤差(較差)を解析すると定誤差, 過失誤差に相当するものがあります。これは計算方法が起因しており, 計算誤差が起きていることがあります。(混合測量に於ける誤差の伝播を参照)

#### 経年変化による誤差

図面成果と実測成果作成時の時間経過の間に観測対象物(境界標等)に動きがあったことが考えられます。観測対象物の動きを経年変化(原始筆界に於ける境界(筆界)の移動を参照)と言いますが一概に経年変化と言っても多種多様でこれを特定しうるのが難しいのです。

境界(筆界)復元に於けるそれぞれの誤差は次の修正或いは除外して使っていきます。計算誤差が見つければ修正して使い, 経年変化は状況によって修正するか, 除くかの判断が必要になってきます。

残ったデータには定誤差, 過失誤差, 復元誤差, 計算誤差, 経年変化の誤差で修正, 除き切れなかったものも含まれます, これらのデータをさらに解析して正規分布から外れるデータは検定によって除外したうえで, さらに残ったデータを解析していく技術が求められます, つまり混合分布が確認される場合は誤差(較差)を要素又は要因別に分けることが必要です。一定方向への定誤差は修正して使うことが必要です。境界(筆界)復元ではデータ数が統計処理に不十分なぐらいに少ない場合があります, この場合は特殊な判断, 解析技術が求められます。

### 重要な誤差の三公理

誤差の現れ方には次の3つの法測があり、誤差の三公理と言います。

第一の公理: 絶対値の小さい誤差は、大きい誤差よりも多く起こる。

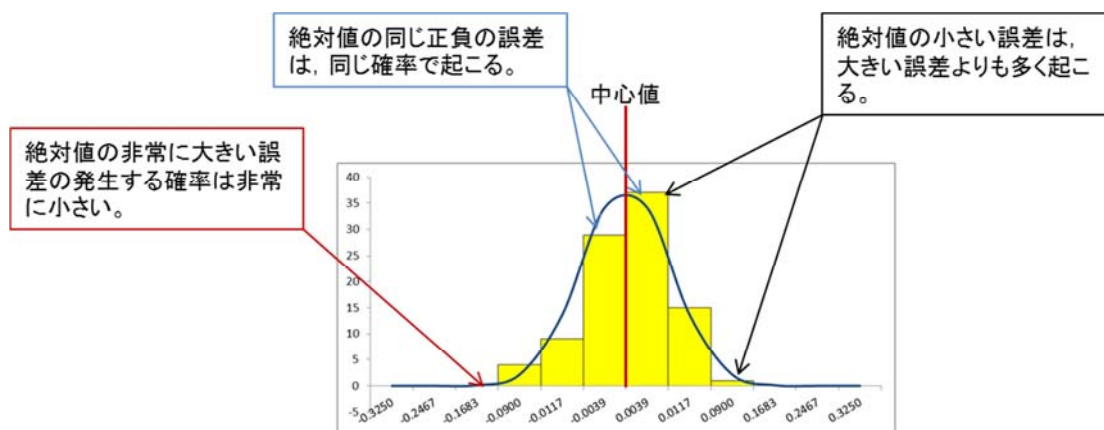
第二の公理: 絶対値の同じ正負の誤差は、同じ確率で起こる。

第三の公理: 絶対値の非常に大きい誤差の発生する確率は非常に小さい。

この三公理は、測定回数が多いときの偶然誤差の起こる確率に特徴的なものです、次の図は誤差をヒストグラムで表したものです、ヒストグラムは一変量のデータの分布を表したグラフです。これで誤差の三公理が説明されます。

青の曲線は正規分布曲線です、データ数が多くなり、誤差の三公理にあてはめれば棒グラフがこの曲線上に載ってきます。

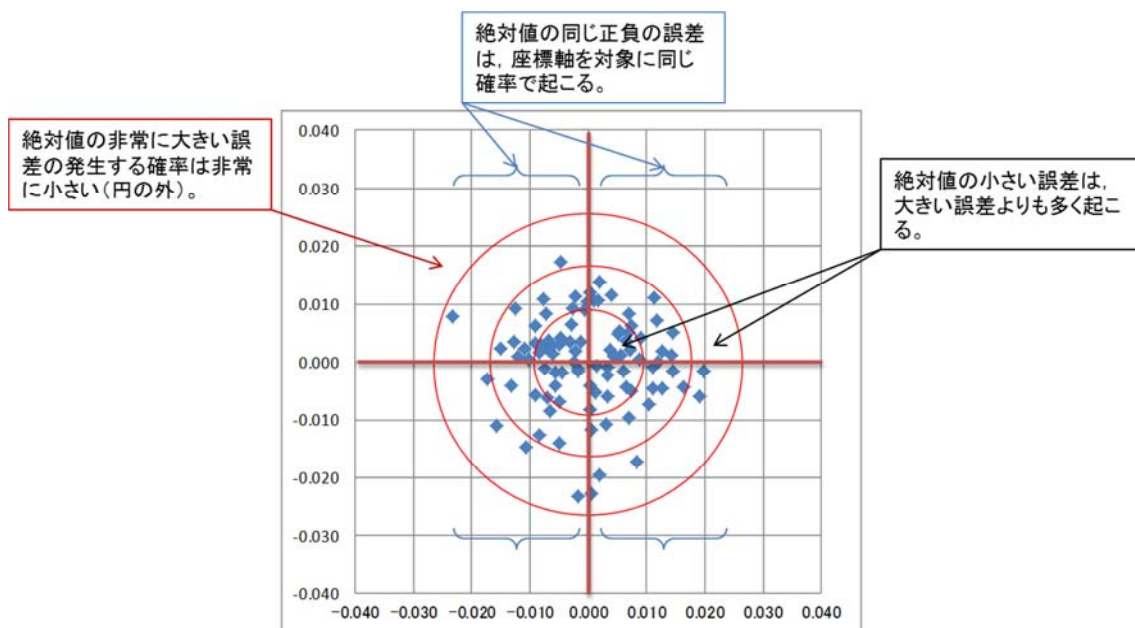
図は実際のデータのグラフです、このように中心値を境に左右対称に近づきます、中心値を境に左側を負、右側を正のデータとすれば中心値に近いほど正、負のデータは多くなり、離れるほど少なくなります。そして正と負の合計は同じ量に近づきます。中心値から極端に離れたデータはほとんどなくなります。



次の図は二変量誤差の分布を散布図に表したものです、境界(筆界)復元では前のヒストグラムよりこちらの散布図の理解が必要です。

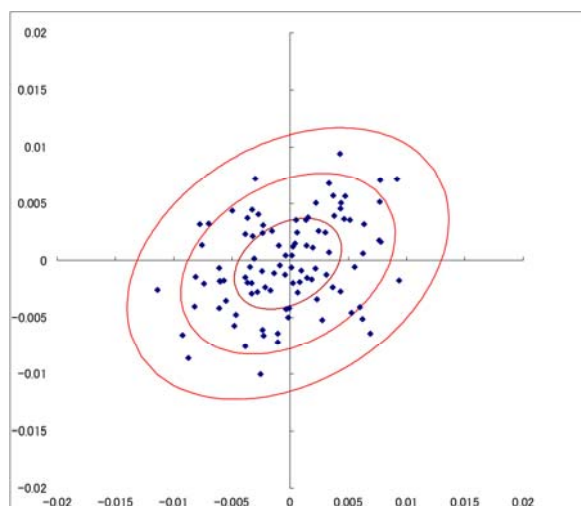
円は内側から1倍標準偏差、2倍標準偏差、3倍標準偏差となっています。二変量誤差は下図のように円の中に収まっている訳ではありませんので複雑になってきます。

誤差が起きる頻度は中心部ほど多く、中心から離れるほど誤差の起きる頻度は少なくなります。



さらに、実際の分布は下図の様に楕円になります、このことがより難しくしている訳です。詳しくは「検定」で触れることにしまして次に進みます。

境界(筆界)復元においてはこの二変量の散布図から正規分布になっているデータを選ぶという作業が必要になります。



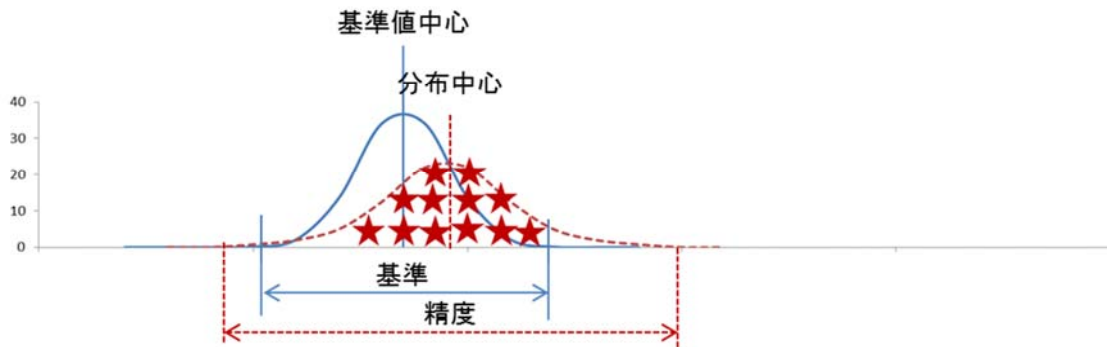
### 正確度と精度

正確度とは、その値が「真値」に近い値であることを示す尺度であり、系統誤差の大小を言う。(確度とも呼ぶ。)精度とは、その複数回の値(複数回の測定または計算の結果)の間での互のばらつきの小ささの尺度であり、偶然誤差の大小を言います。正確度は正確性の尺度であり、精度は再現性の尺度です。

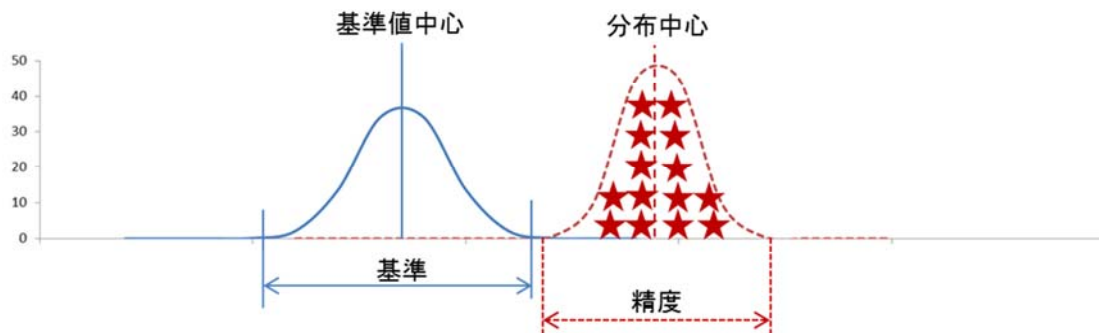
## 一変量の分布

一変量の分布の高精度、低正確度の状態をグラフで表すと下図のようになります。

高正確度だが低精度



高精度だが低正確度



このどちらが正確はこの図だけでは判りませんが、「公差と不良率」で説明しますが簡単に言えば不良率で考えれば単純に判ります。

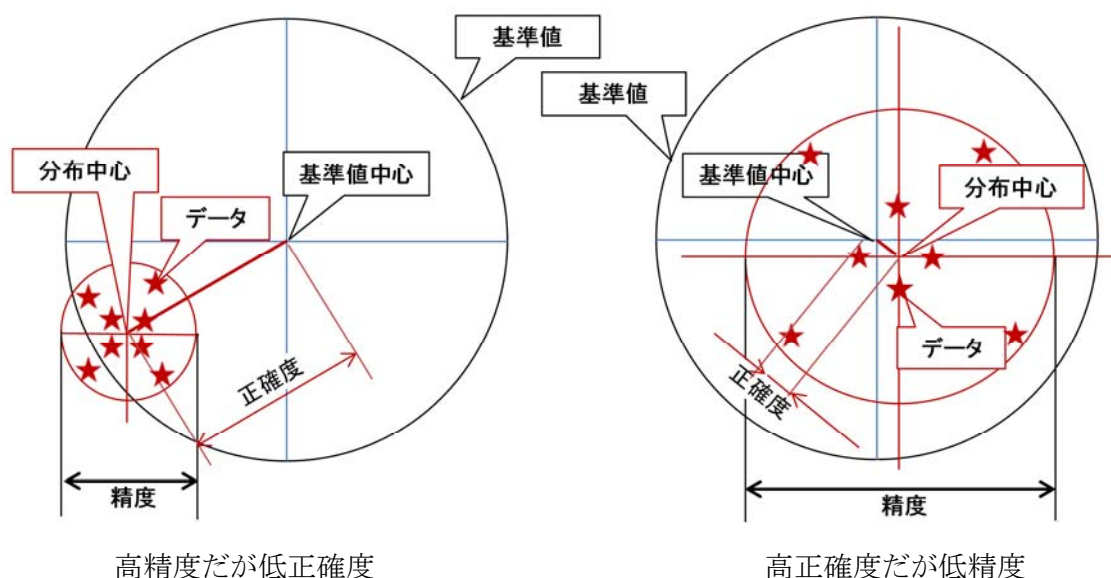
公差とか基準が無いときには仮定の公差を当て嵌めて比較すれば良いのです。これは次に説明する二変量についても同じです。

これらの違いを説明すると的の中心を基準値中心として、ここでは繰り返し測定することを的に向かって矢を何度も放つことになぞらえると正確度とは、的の中心と矢が当たった場所との近さを意味します。中心に近ければ近いほど正確度は高いと見なされる。矢をどんどん放ったとき、矢の当たった範囲の大きさが精度に相当する。矢を1回だけ放った場合、条件が同じであれば以前に得られた精度の範囲でその後の矢が似たような場所に当たると予想できます。

全ての矢が非常に狭い範囲に当たった場合、精度が高いということができ、これはその範囲が中心とどういう距離にあるのかは無関係です。精度の高い測定は必ずしも正確度が高いとは限りません。



## 二変量の分布(座標値)



### 精度の定量化

精度は測定における標準偏差で表されることが多く、これを測定の標準誤差と呼ぶこともある。標準偏差を正確に求めるために十分な回数の測定が行われ、測定の誤差に偏りがなければ、測定のうち、一変量では68.3%は  $1\sigma$  の範囲にあり、65.44%は  $2\sigma$  の範囲にあり、99.7%は  $3\sigma$  の範囲にあることになる。

これは測定を繰り返して平均を求める場合にも適用できる。この場合は特に標準誤差という用語が使われる。

### 正確度の定量化

正確度に関しては、次のことを区別することができる、測定値の平均と標準値の差、すなわち偏り。この値を確定して補正することは校正において必須とされる。

精度と正確度を統合して判断する規準が測量業界にはないが、今まで使われて平均二乗誤差が適当ではないかと考えられる。(標準偏差と平均二乗誤差の違いを参照)

公差、最悪の場合この範囲に観測値又は計算値がなければならないとする基準値を公差と言う、この公差の定めがある場合は公差に対する不良率を推定しその成果全体の良否を判断することになります。(公差と不良率を参照)

### 境界(筆界)復元に於ける正確度について

境界(筆界)復元においては主に図面の値を実測の値に変換して使うので正確度については問題にしない、変換は図形の方位の違いを一致させる回転と図形の大きさを合わせる伸縮

を行う、回転については図面が作られた時に図形の方角を何処にするかが明確でなかったとか、磁北を基準にする場合でもその精度が確保出来なかったので図形の方角は一致しない、伸縮についてはその時々に使われた距離測定器が異なる、末尾の処理が異なるなどの要因が重なるのである程度の範囲ではやむを得ない伸縮が起きている、公図(土地台帳附属地図)を図面值として変換を行う場合に、図面対実測の伸縮率の大きさが問題視されることがあるのでこのような場合はその大きさについてその要素の説明が出来なければなりません。

境界(筆界)復元で扱うデータは距離だけ、角度だけというデータの種別が一書類だけのデータを一変量のデータと言います。

それに対して、座標値の場合X座標値とY座標値で一つの点を示すデータを二変量のデータと言います。

境界(筆界)復元ではこの二つの誤差について理解する必要があります、一般的には全て一変量で説明されますがこれでは不十分です。