日本分析化学会第77回分析化討論会

公開シンポジウム：　分析化学における実験データの正しい扱い方

なぜ実験データの解析は大事か？

一般公開

2017/05/27

加納健司（京大院農）

前田耕治（京工繊大院工）

高精度測定から生まれたパラダイムシフト

円形運動から楕円運動

Tycho Brahe
平均で1分，最高で2秒の観測精度
(1分 = 1/60º　1秒 = 1/3600º)

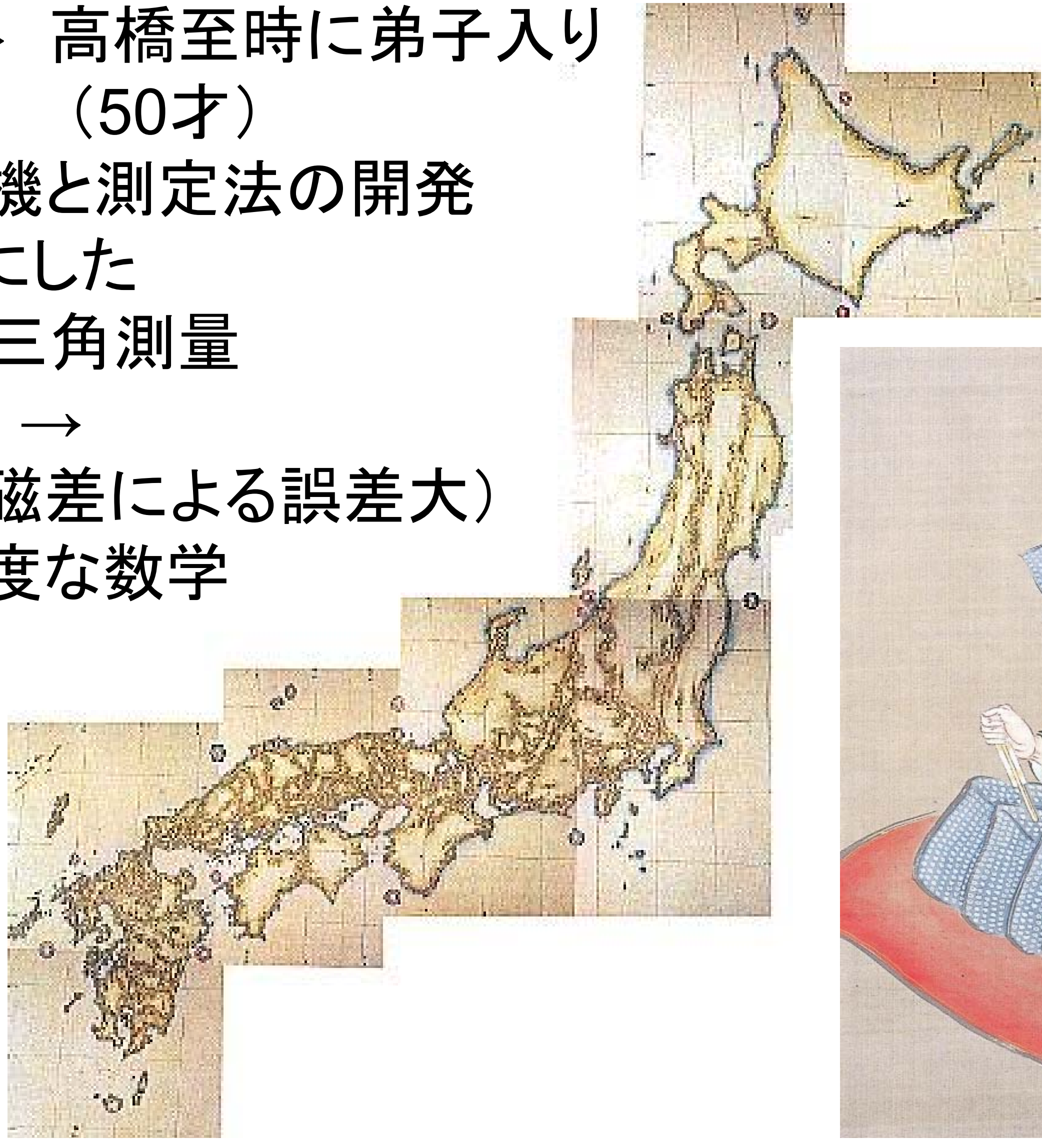
Johannes Kepler
計算値の8分のずれをも許さず，新しい法則を提唱

そして，
ガリレオの天動説が生まれ，
ニュートンの運動力学の成功につながる

伊能忠敬の高精度測量

日本地図の南北の誤差は最大でも1/1000

造り酒屋名　→　高橋至時に弟子入り
（50才）
独自の測量器機と測定法の開発
北極星を目標にした
大きな三角測量
（西洋:磁石 →
磁差による誤差大）
江戸時代の高度な数学



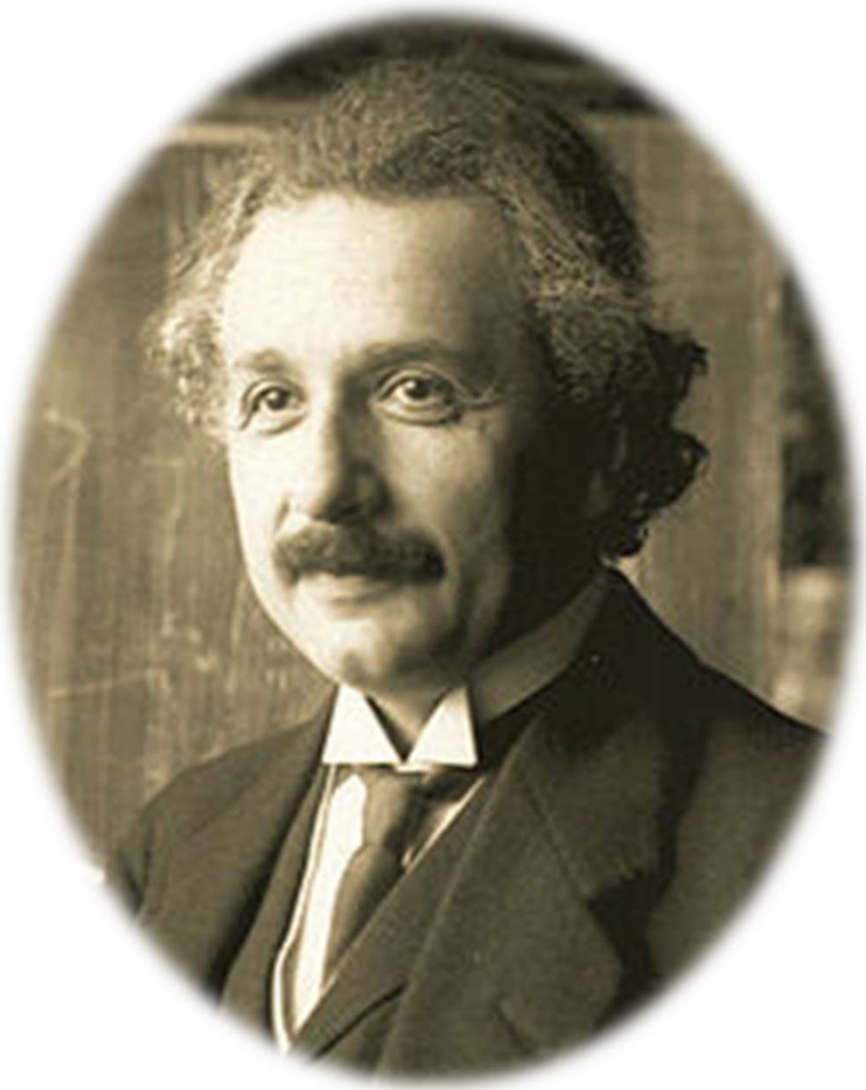
Random Walk
酔歩

1905

奇跡の物理の年

光電効果の理論
ブラウン運動の理論
特殊相対性理論

原子・分子の存在の実証

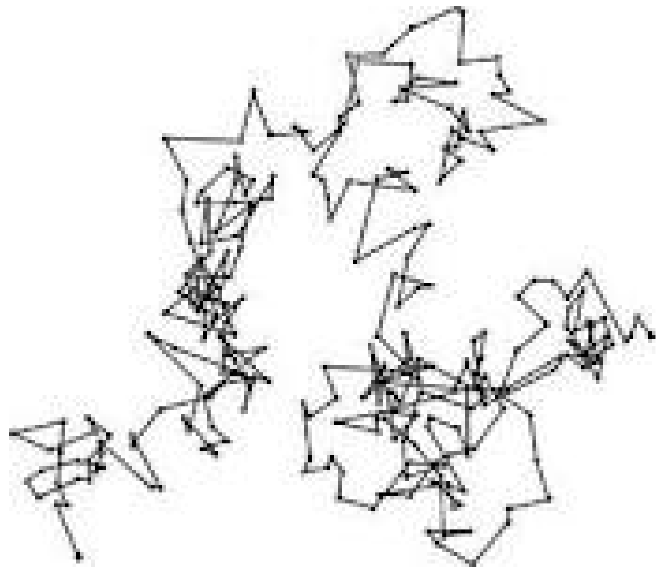


Albert Einstein 1879-1955

ブラウン運動 (Brownian Motion)

1827

Robert Brown (1773 – 1858)



水中で花粉が動く!?

1910年　長岡半太郎　（原子模型）

花粉の大きさは通常、直径30 ― 50 μm
水分子　0.3 nm！

１円玉でタンカーを動かす！？

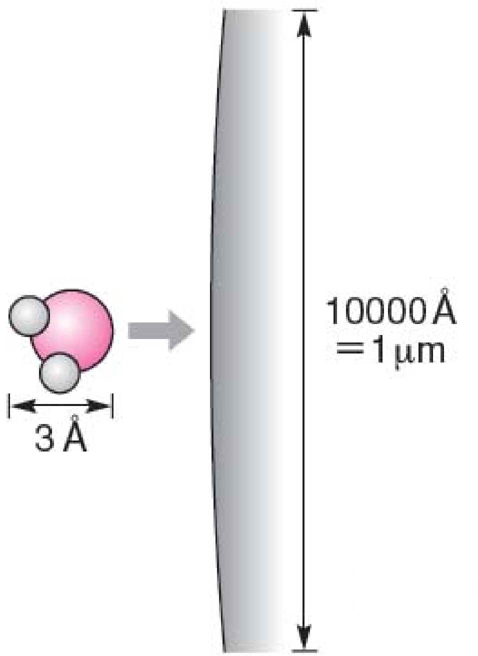
体積比を考え，それを重量比にすると

H_2O　　$\phi = 3$ Å

ポリスチレン微粒子
　　$\phi = 1$ μm　(= 10^4 Å)

1円玉　1 g

石油タンカー
　　20 万トン＝ 2×10^{11} g



水分子とブラウン運動を
する粒子の大きさの比較

原子論にまつわる100年に亘る論争の終止符

アインシュタインの式

$$\sqrt{<x^2>} = \sqrt{2Dt} \quad (\Delta x = \sqrt{2D\Delta t})$$

ストークス－アインシュタインの式

$$D = \frac{RT}{6\pi\eta r N_A}$$

アボガドロ数の実験的算出

$$N_A = (6.5 - 6.9) \times 10^{23}\ \mathrm{mol}^{-1}$$

$$N_A = 6.02 \times 10^{23}\ \mathrm{mol}^{-1}$$

1908

Jean Baptiste Perrin, 1870–1942

自然科学、特に医学生物学分野での論文の約**70**%が再現不能であることがわかった！

BIOMEDICINE

NIH mulls rules for validating key results

US biomedical agency could enlist independent labs for verification.

BY MEREDITH WADMAN

In biomedical science, at least one thing is apparently reproducible: a steady stream of studies that show the irreproducibility of many important experiments.

In a 2011 internal survey, pharmaceutical firm Bayer HealthCare of Leverkusen, Germany, was unable to validate the relevant preclinical research for almost two-thirds of 67 in-house projects. Then, in 2012, scientists at Amgen, a drug company based in Thousand Oaks, California, reported their failure to replicate 89% of the findings from 53 landmark cancer papers. And in a study published in May, more than half of the respondents to a survey at the MD Anderson Cancer Center in Houston, Texas, reported failing at least once in attempts at reproducing published data (see 'Make believe').

The growing problem is threatening the reputation of the US National Institutes of Health (NIH) based in Bethesda, Maryland, which funds many of the studies in question. Senior NIH officials are now considering adding requirements to grant applications to make experimental validations routine for certain types of science, such as the foundational work that leads to costly clinical trials. As the NIH pursues such top-down changes, one company is taking a bottom-up approach, targeting scientists directly to see if they are willing to verify their experiments.

There is the looming

➲ NATURE.COM
For more on the challenges of reproducibility:
go.nature.com/zqtrnp

「統計手法」を明確にするよう*Nature*編集方針を変更

*P*値の誤解，誤用，不正使用の蔓延に米国統計学会が警告

R. L. Wasserstein, N. A. Lazer, *Am. Stat.*, **70**, 129 (2016)
M. Baker, D. Penny, *Nature*, **531**, 151 (2016)

統計リテラシーの欠如

諏訪雅頼, ぶんせき, **2017**, 122

不正（疑惑）の例

メンデルの法則（？）

ベル研究所
ソウル大学
某製薬会社
理化学研究所

では，現実的な事象として，
以下の問に答えられますか

117.0 ± 2.2 と書いたら,
有効数字は何桁？

正しい有効数字の表記は？

正しいエラーバー表記は？

アナログ目盛りは何桁まで読み取る？

デジタル表示データの有効数字は？

許容誤差 0.02 cm^3 のホールピペットで10 cm^{3}を 5回 測り取ったら、標準偏差が±0.01 cm^3 だった。この誤差はどう表せばいい？

系統誤差と偶然誤差の統一的表記

→　「不確かさ」とは？

有効数字を考慮した誤差の計算方法は？

初等教育で習った加減乗除の桁あわせ法では対処できない！

非線形な関係を対数変換や逆数変換した場合の誤差は？

対数変換による誤差

$$c = c_0 \exp(-kt) \quad \Rightarrow \quad \ln(c/c^{\ominus}) = -kt + \ln(c_0/c^{\ominus})$$

逆数変換と傾き・切片からの物理量評価に関わる誤差

$$v = \frac{V_{\max}}{1 + K_{\mathrm{M}}/c_{\mathrm{S}}} \quad \Rightarrow \quad \frac{1}{v} = \left(\frac{K_{\mathrm{M}}}{V_{\max}}\right)\frac{1}{c_{\mathrm{S}}} + \frac{1}{V_{\max}}$$

Michelis-Menten式　　Lineweaver-Bulkプロット

吸光度測定の誤差

$$A \equiv -\log\left(\frac{I}{I_0}\right) = -\log t = \varepsilon c l$$

どの検定法を使うの？

類似の試料AとBのある成分についてそれぞれN_A回，N_B回分析したとき，試料Aと試料Bは同一のものかどうか判断するにはどうする？

同一試料のある成分についてAさんとBさんがそれぞれN_A回，N_B回分析したとき，AさんとBさんのデータ群に差があるかどうかを知りたいときはどうする？

その差がないと判断できたとき，２人のデータの違いをどう表現する？（統計的には他方のデータだけを採用することはできない）

京都市下水道法では、「アルキル水銀はN.D.（検出されない）」とされている。

「検出されない」ことの基準は？

このセンサーは5.0 μM まで測れます！

「検出限界」と「定量限界」の違いは？

100枚の画像の中で１枚だけ，“都合の良い”画像があった場合，その画像を示すことで，実験がうまくできたことの証明になるか？

画像処理は捏造の温床・・・

いくつかの問に対して
答えられなかった課題を自覚して，
各講演者の話を聴いてください！

実験データの有効数字を理解する

（京大院農）北隅 優希

検出限界と定量下限、信頼性に関わる用語

（明星大院理工）上本 道久

検定の考え方と実際

（京大院工）西 直哉

最小二乗法と検量線

（阪大院理）諏訪 雅頼

不確かさの概念と無機分析における見積もり例

（明星大院理工）上本 道久

ランダムウォーク（二項分布）から誤差論（正規分布）へ

$$G_{\mu,\sigma}(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right)$$

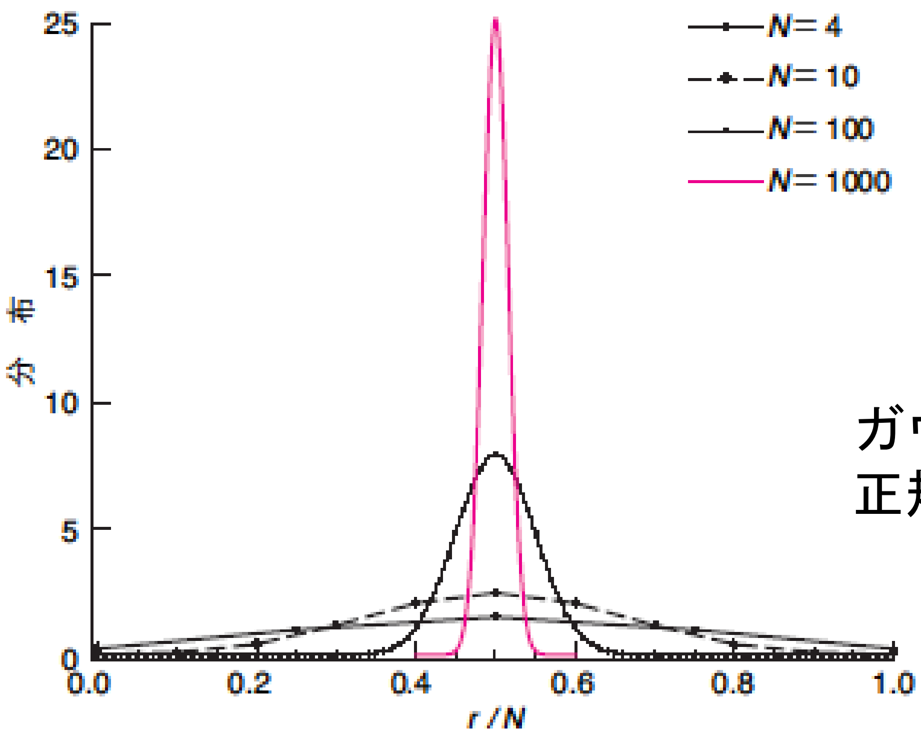


図 2　右方向に運動する分子の確率の N 依存性

標準偏差　σ
分散　σ^2

ガウス分布　**(Gaussian distribution)**
正規分布**(normal distribution)**

正規分布

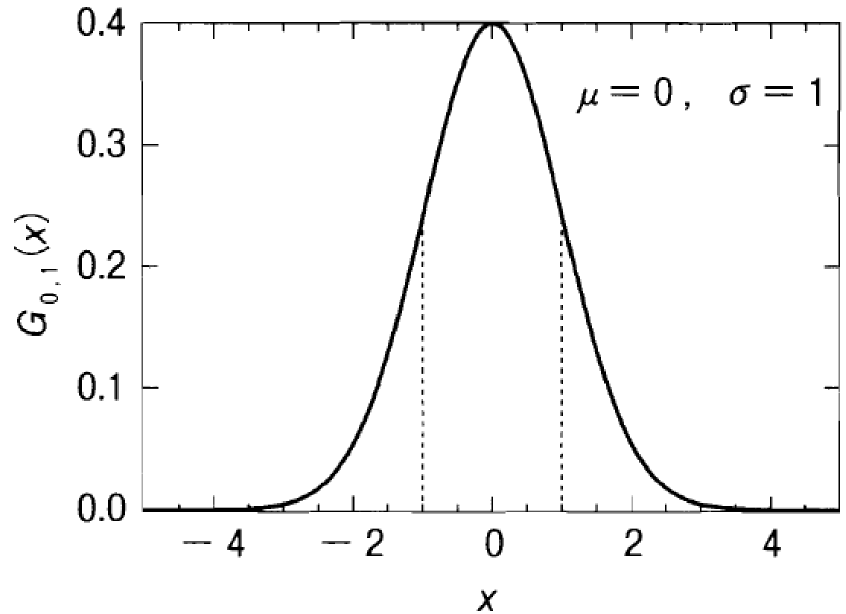
$$G_{\mu,\sigma}(x)=\frac{1}{\sqrt{2\pi\sigma}}\exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right)$$

μ: 母平均, σ: 母標準偏差, σ^2: 母分散

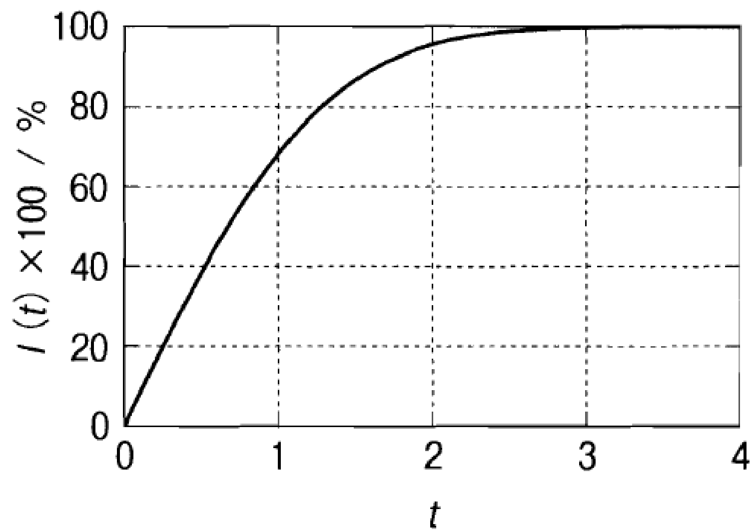
平均値から±$t\sigma$の確率

t	%
1	68.3
2	95.4
3	99.7
4	99.99
5	99.9999

(a)

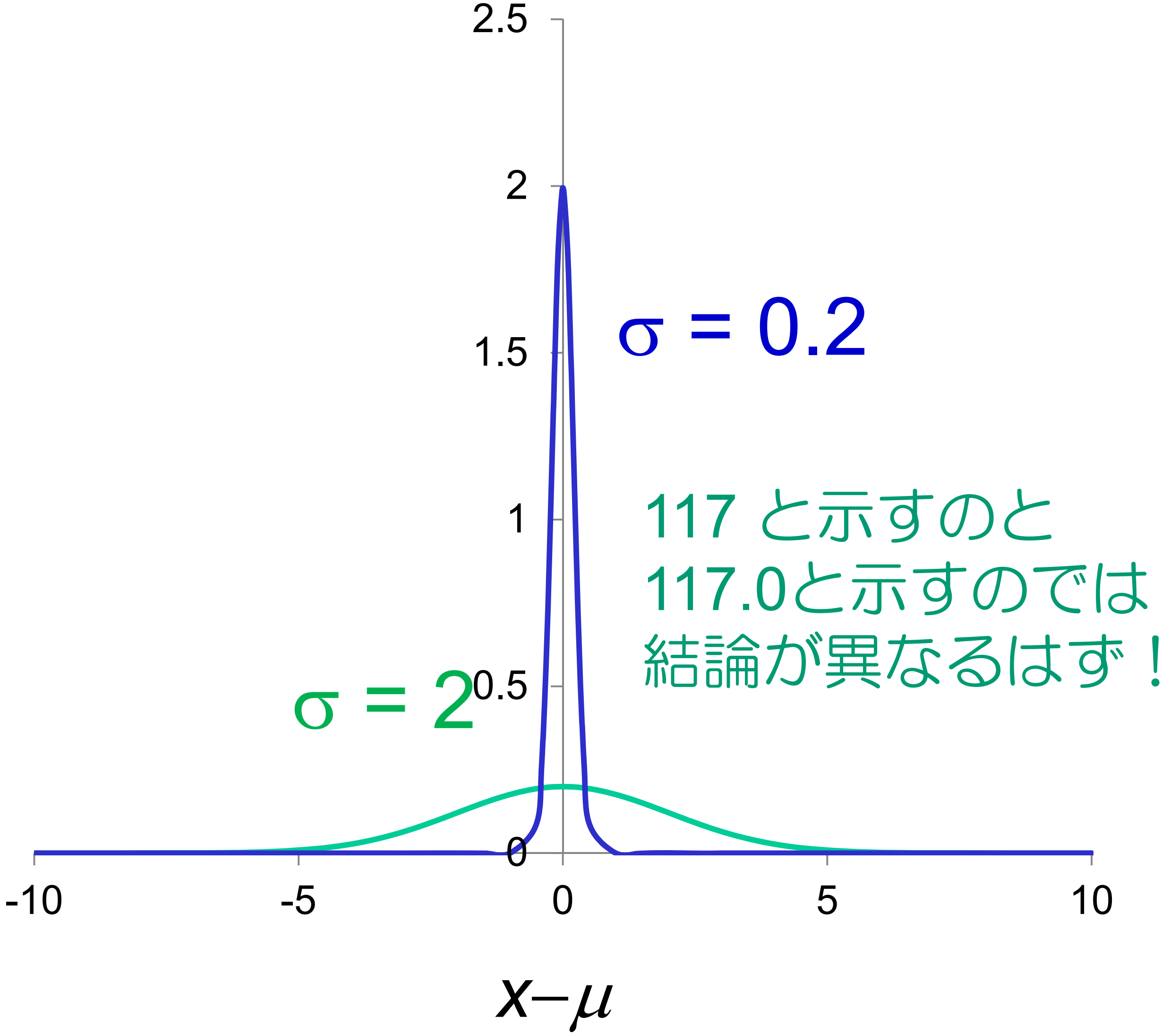


(b)

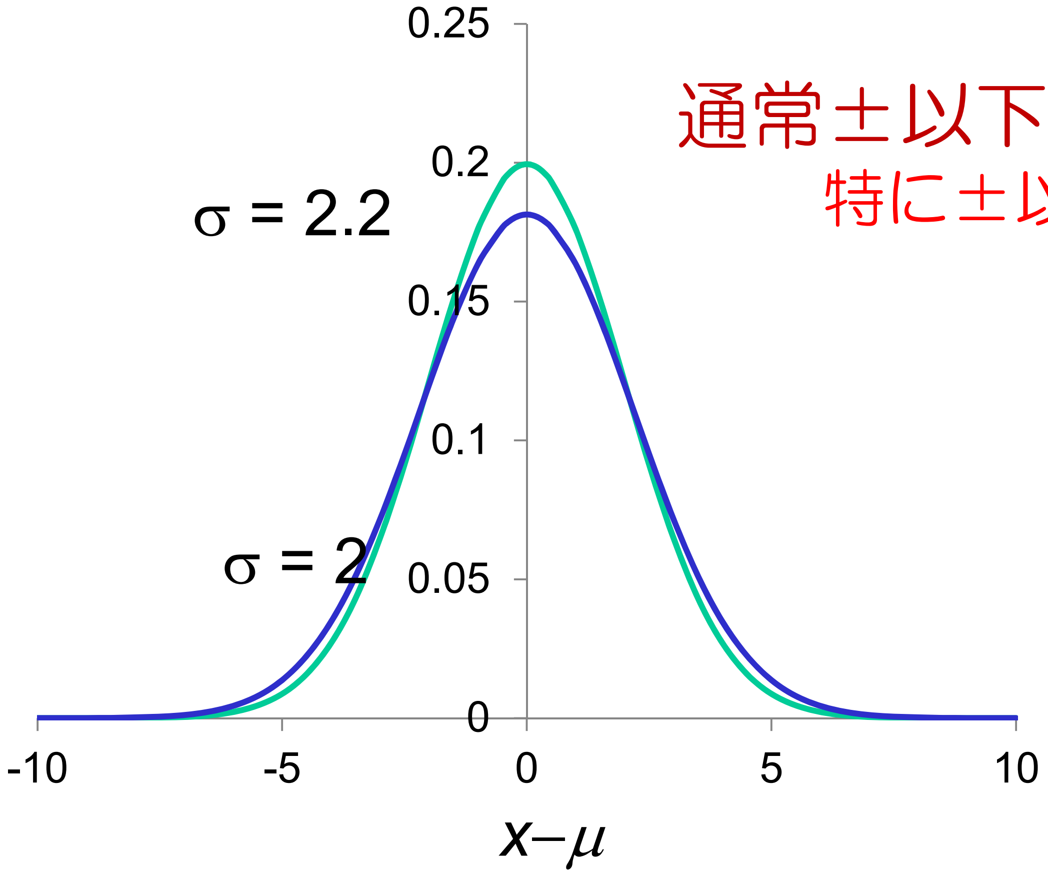


ある人の話:「◯×さんは，3σだ！ いや5σかも･･･!?」 この意味は？
ついでに，あの“偏差値”って何？

117 ± 2 と 117.0 ± 0.2の違いは？



117.0 ± 2.2 と117.0 ± 2の違いは



実質同じ（117.0 ± 2.2 は実質上有効数字3ケタ）

通常±以下は１ケタ表記で十分

特に±以下の数字が9に近い場合！

自然科学系の実験値の誤差表示はどれが適切？

母標準偏差

$N_{\mathrm{population}}$の表示

(標本)標準偏差

N_{sample}の表示

(N_{sample}が同じとき比較可能)

信頼限界

Zあるいはpの表示

(N_{sample}が異なっても比較可能)

自然科学ではほとんどの場合、信頼限界をつかうべきなのですが・・・

信頼区間

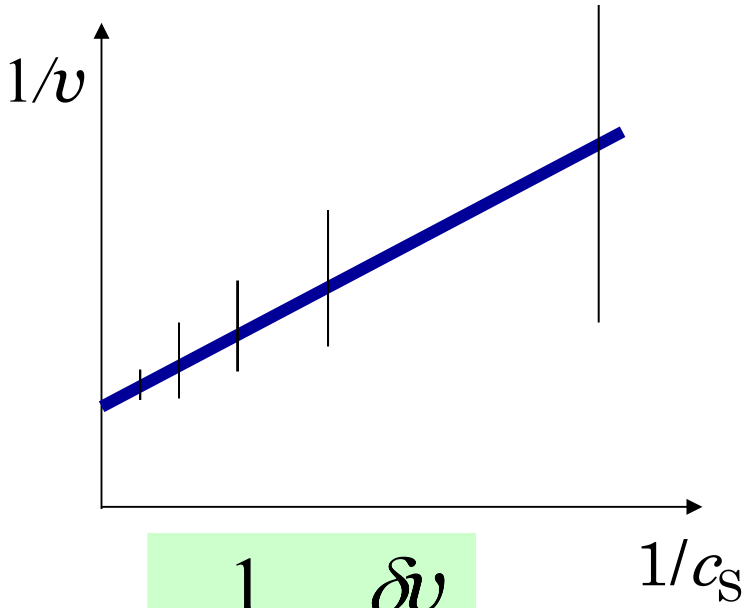
117.6±0.2の意味は？

信頼度 z(%) = 95%

逆数プロットの線形最小二乗法
Lineweaver-Bulkを例に

$$\mathrm{E} + \mathrm{S} \rightleftarrows \mathrm{ES} \longrightarrow \mathrm{E} + \mathrm{P}$$

$$v = \frac{V_{\mathrm{max}}}{1 + K_{\mathrm{M}} / c_{\mathrm{S}}}$$



$$\delta \frac{1}{v} = \frac{\delta v}{v^2}$$

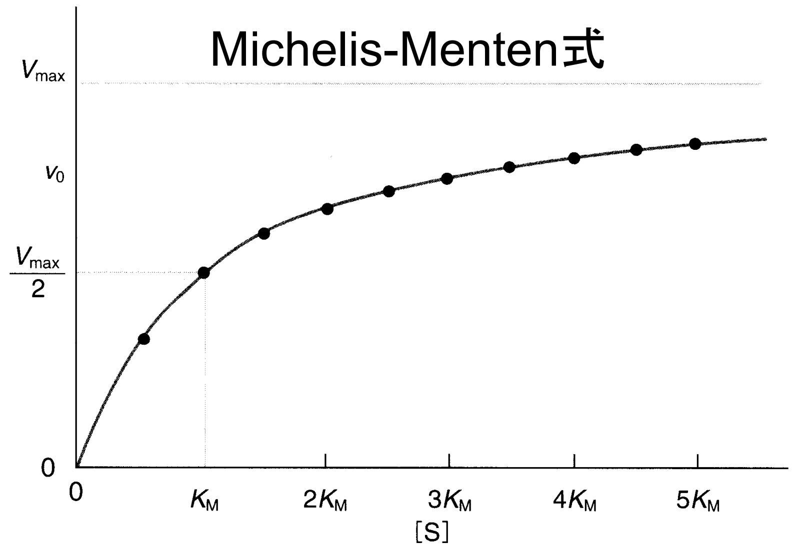


図 14・8　簡単なミカエリス・メンテン反応の初速度 v_0 と [S] の関係　基質濃度を 0.5 K_{M} おきに 0.5 K_{M} から 5 K_{M} までプロット．
Animated Figures

$$\frac{1}{v} = \left(\frac{K_{\mathrm{M}}}{V_{\mathrm{max}}}\right)\frac{1}{c_{\mathrm{S}}} + \frac{1}{V_{\mathrm{max}}}$$

試料濃度を明確にすべき溶液の調整・希釈法は？

実験データの有効数字を理解する
(分析化学における実験データの正しい扱い方)

北隅優希
京都大学大学院農学研究科
kitazumi.yuki.7u@kyoto-u.ac.jp

第77回分析化学討論会
2017年5月27日
龍谷大学深草キャンパス

物理量の取り扱い

先生「この溶液の濃度を教えて？」

A君「5．3261938です」

先生「桁数多いな～」

A君「電卓で計算したので間違ってませんよ」

先生「そうではなくて，それと単位はmMでいいのか？」

問題点1　有効数字(誤差)について考慮されていない

問題点2　単位が不明確

データ取扱いの基礎
ー測定と誤差ー

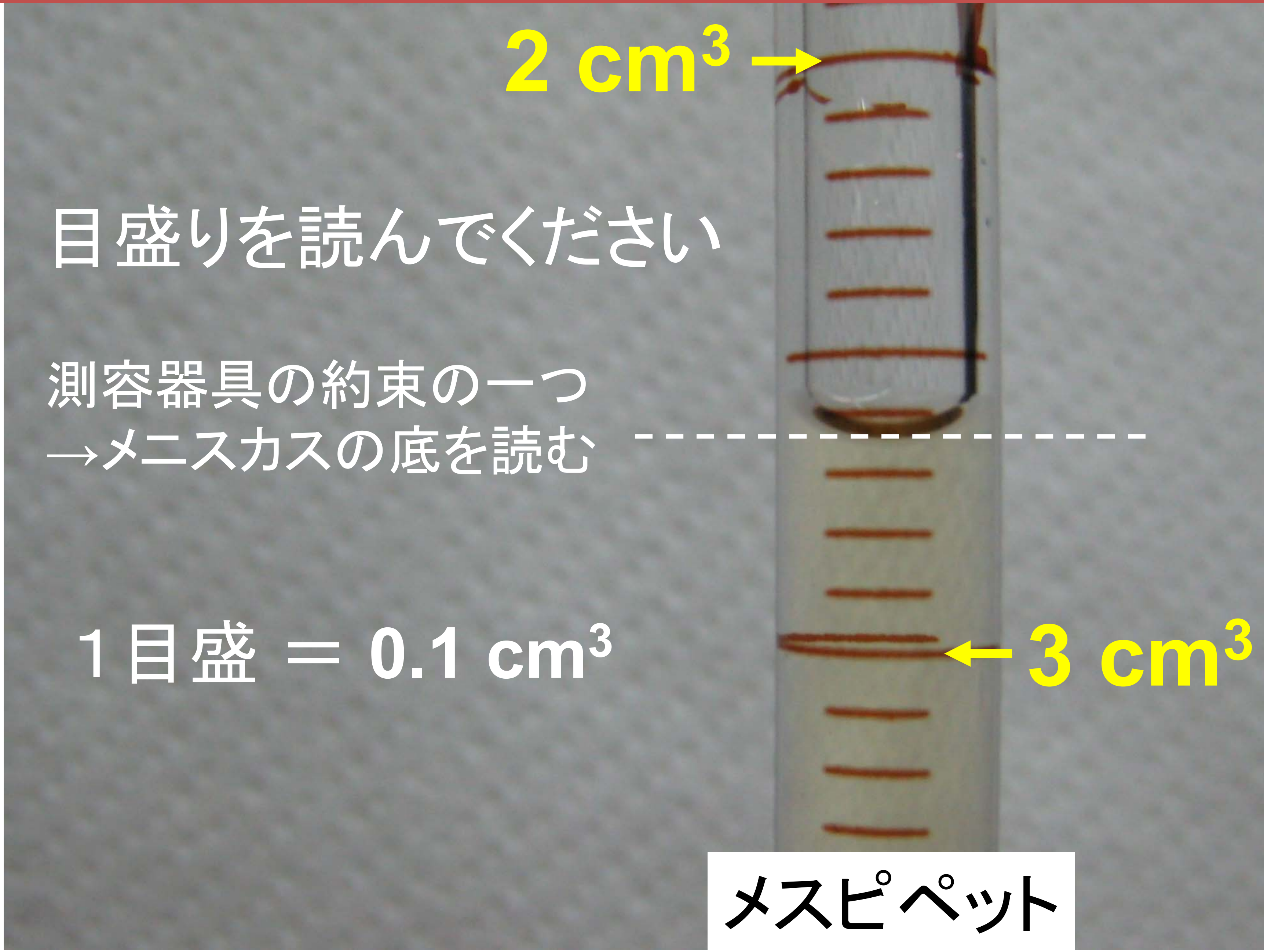
- 誤差　測定値から真の値を引いた差
- 真の値　実際には求められない観念的な値
- 有効数字

　誤差を含んだ測定値の最も単純な表し方
　測定結果などを表す数字のうちで，位取りを示すだけのゼロを除いた意味のある数字
- 誤差伝播

　データ処理にともなう誤差の伝わり方
- 信頼区間

　ある確率(信頼度)で真の値を含む範囲

アナログ目盛りの読み取り



目盛りを読んでください

測容器具の約束の一つ
→メニスカスの底を読む

1目盛 = 0.1 cm^3

どのように読みましたか？

1目盛 ＝ **0.1 cm^3**

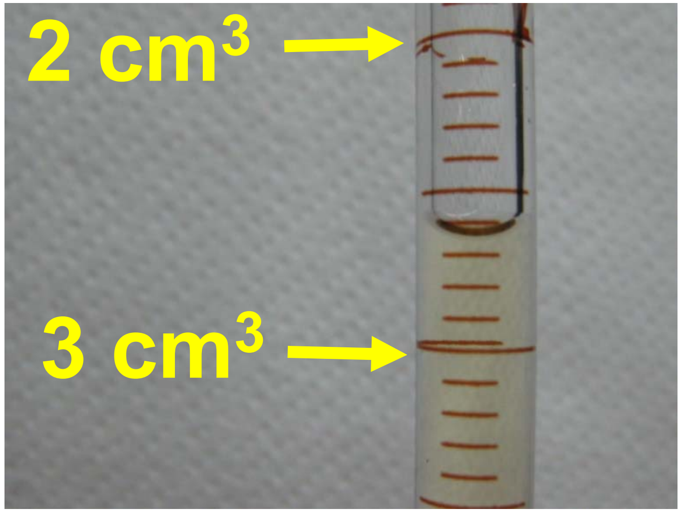
× 2.6　cm^3

○ 2.62

○ 2.63

○ 2.64

× 2.7



原則：アナログ目盛は最小目盛の1/10まで読む

有効数字とは

有効数字　＝　測定結果をその測定精度と合わせて
表示するために必要な数字の桁
⇒確定的な全ての桁（●）＋不確かな一桁（×）

最小目盛 が0.1 cm^{3}なら
その1/10の0.01 cm^{3}の桁が不確か

●　●　×
2.63 cm^3

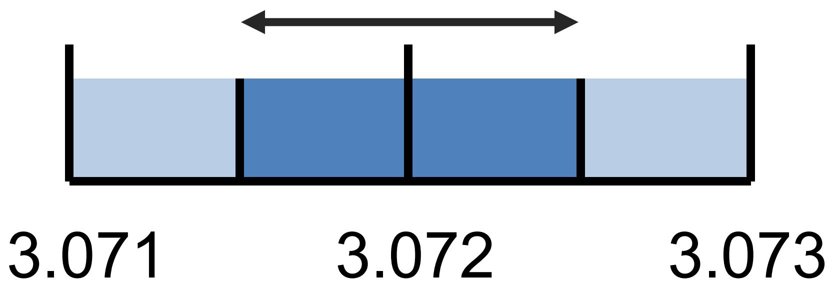
<u>測定値とその精度を同時に表現</u>

有効数字のその最終桁には±1の誤差が含まれる(だろう)
⇒2.63 ± 0.01 cm^3

デジタル表示の有効数字

デジタル表示

3.072



3.0720±0.005 を
意味する場合もある

もしも表示が
ふらついていたら？

3.144

3.039

3.128

3.054

3.091

上記の例なら
3.1±0.1　が妥当

デジタル表示器

装置の性能も確認すべき

例　電圧計
表示　1.386942 V で安定

装置の仕様(例)

分解能　1 μV
確度　0.025% + 40 μV

測定の精度

1.386942 V の 0.025% すなわち　0.00035 V　に 40 μV を加えた
0.00039 V　以内であることを意味する

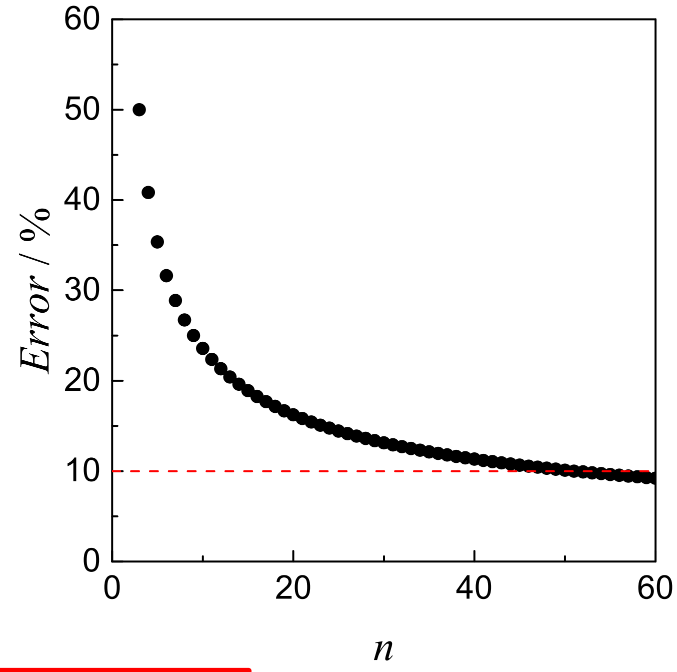
測定結果の適切な表現は
(1.3869 ± 0.0004) V

誤差の誤差

誤差も測定結果から評価される ⇐ 当然ばらつく

- 誤差に分布がある

誤差の相対誤差 : $\dfrac{1}{\sqrt{2n-2}}$　n…測定回数



二桁の誤差を主張するためには…

誤差1を1.0と主張するには
⇒誤差の相対誤差10%以下
⇒測定数51回以上！

かなり多い測定が必要となる

原則 : 誤差は一桁に丸める

桁数のルール

装置で測定値が105.4236と表示された。適切な表現は

誤差が0.2なら

105.4~~236~~±0.2 → 105.4±0.2

誤差が2なら

105~~.4236~~±2 → 105±2

原則：値の最終桁は誤差と同じにする

有効数字の正しい表し方

1000 有効数字を明確に表現するなら

		有効数字
1	$\times\ 10^3$	（1桁）
1.0	$\times\ 10^3$	（2桁）
1.00	$\times\ 10^3$	（3桁）
1.000	$\times\ 10^3$	（4桁）
仮数部	指数部	

例

(386 ± 4) kg ならば

$(3.86 \pm 0.04) \times 10^2$ kg　と表記するのが最も好ましい

誤差はどのように伝わる？誤差伝播

それぞれの誤差と、計算結果の関係

関数

$$q(x, y, z, ..)$$

誤差を持つ測定値

$$x + \delta x$$

$$y + \delta y$$

$$z + \delta z$$

$$..$$

qの誤差δqは、

$$\delta q = \sqrt{\left(\frac{\partial q}{\partial x}\delta x\right)^2 + \left(\frac{\partial q}{\partial y}\delta y\right)^2 + \left(\frac{\partial q}{\partial z}\delta z\right)^2 \cdots}$$

条件
- 測定値は互いに独立
- 誤差は偶然誤差のみ

誤差伝播の式の導出

$q = f(x_1, x_2, \ldots, x_n)$

qは任意の分布を持つ　平均q_0　分散δq^2

x_iも任意の分布を持つ　平均x_{i0}　分散δx_i^2

qを1次までTaylor展開

$$q = q_0 + \sum_i (x_i - x_{i0}) \left(\frac{\partial q}{\partial x_i} \right)_{x_i - x_{i0}}$$

$$q - q_0 = \sum_i (x_i - x_{i0}) q_{xi}$$

where

$$q_{xi} = \left(\frac{\partial q}{\partial x_i} \right)_{x_i = x_{i0}}$$

$$\begin{aligned}
\delta q^2 &= \mathrm{E}\left[(q - \mathrm{E}[q])^2 \right] \\
&= \mathrm{E}\left[(q - q_0)^2 \right] \\
&= \mathrm{E}\left[\left(\sum_i (x_i - x_{i0}) q_{xi} \right)^2 \right] \\
&= \mathrm{E}\left[\sum_i (x_i - x_{i0})^2 q_{xi}^2 \right. \\
&\qquad \left. + \sum_{i \neq j} (x_i - x_{i0})(x_j - x_{j0}) q_{xi} q_{xj} \right] \\
&= \mathrm{E}\left[\sum_i (x_i - x_{i0})^2 q_{xi}^2 \right] \\
&= \sum_i q_{xi}^2 \mathrm{E}\left[(x_i - x_{i0})^2 \right] \\
&= \sum_i q_{xi}^2 \delta x_i^2
\end{aligned}$$

（Eは期待値、
qの期待値はq_0）

（共分散はゼロ）

$$\boxed{\delta q = \sqrt{\left(\frac{\partial q}{\partial x} \delta x \right)^2 + \left(\frac{\partial q}{\partial y} \delta y \right)^2 + \left(\frac{\partial q}{\partial z} \delta z \right)^2 \cdots}}$$

しばしば使うふたつの形式

$$\delta q = \sqrt{\left(\frac{\partial q}{\partial x}\delta x\right)^2 + \left(\frac{\partial q}{\partial y}\delta y\right)^2 + \left(\frac{\partial q}{\partial z}\delta z\right)^2 \cdots}$$

一般式が苦手な方へ、よく使う便利な形を紹介

加減算

$$q = x + y - z \cdots$$

絶対誤差の二乗和

$$\delta q^2 = \delta x^2 + \delta y^2 + \delta z^2 \cdots$$

乗除算

$$q = \frac{xy \cdots}{z \cdots}$$

相対誤差の二乗和

$$\left(\frac{\delta q}{q}\right)^2 = \left(\frac{\delta x}{x}\right)^2 + \left(\frac{\delta y}{y}\right)^2 + \left(\frac{\delta z}{z}\right)^2 \cdots$$

加減算の誤差伝播

$$q \pm \delta q = (x \pm \delta x) + (y \pm \delta y)$$

原則：加減算の誤差は絶対誤差の二乗和で伝播

$$(\delta q)^2 = (\delta x)^2 + (\delta y)^2$$

$$\delta q = \sqrt{(\delta x)^2 + (\delta y)^2}$$

例

$(2.3 \pm 0.1) + (3.32 \pm 0.01)$

$\sqrt{0.1^2 + 0.01^2} = 0.10\cdots$ なので

$= 5.6 \pm 0.1$

加減算の誤差は測定値の中における最大の絶対誤差で決まる

有効数字と加減算

簡略化すると、有効数字の加減算の桁合わせになる

$$10.57 + 2.345 + 120.4 =$$

$$\begin{array}{r} 10.57 \\ 2.345 \\ +\quad 120.4 \\ \hline 133.315 \end{array}$$

最も高い最終桁
最大の絶対誤差

$$\cong 133.3 = 1.333 \times 10^2$$

有効数字の原則：加減算では最も高い最終桁に合わせる

乗除算の誤差伝播

$$q \pm \delta q = \frac{(x \pm \delta x)(y \pm \delta y)}{(z \pm \delta z)}$$

原則:乗除算の誤差は相対誤差の二乗和で伝播

$$\left(\frac{\delta q}{q}\right)^2 = \left(\frac{\delta x}{x}\right)^2 + \left(\frac{\delta y}{y}\right)^2 + \left(\frac{\delta z}{z}\right)^2$$

$$\delta q = q\sqrt{\left(\frac{\delta x}{x}\right)^2 + \left(\frac{\delta y}{y}\right)^2 + \left(\frac{\delta z}{z}\right)^2}$$

例

$\dfrac{(39\pm 1)(6\pm 1)}{(310\pm 1)}$の場合

δqは

$$0.75\sqrt{\left(\frac{1}{39}\right)^2 + \left(\frac{1}{6}\right)^2 + \left(\frac{1}{310}\right)^2}$$

$$= 0.12$$

よって

$$0.7 \pm 0.1$$

乗除算の誤差は測定値の中における最大の相対誤差で決まる

有効数字と乗除算

簡略化すると、有効数字の乗除算における桁合わせになる

5桁 3桁 4桁

$$1.2345 \times 15.4 / 238.7$$

$$= 0.079645161$$

$$\approx 0.0796$$ 3桁

15.4が3桁で桁数最小(相対誤差最大)
演算後:4桁目を四捨五入して、3桁にする

有効数字の原則:乗除算では最小の桁数に合わせる

複雑な関数の誤差伝播の例

$$q = \ln \frac{y}{y^{\ominus}}$$

$$q \pm \delta q = \ln \frac{(y \pm \delta y)}{y^{\ominus}}$$

$$\delta q = \sqrt{\left(\frac{\delta y}{y}\right)^2} = \frac{\delta y}{y}$$

対数関数

$$q = A\exp(-Bt)$$

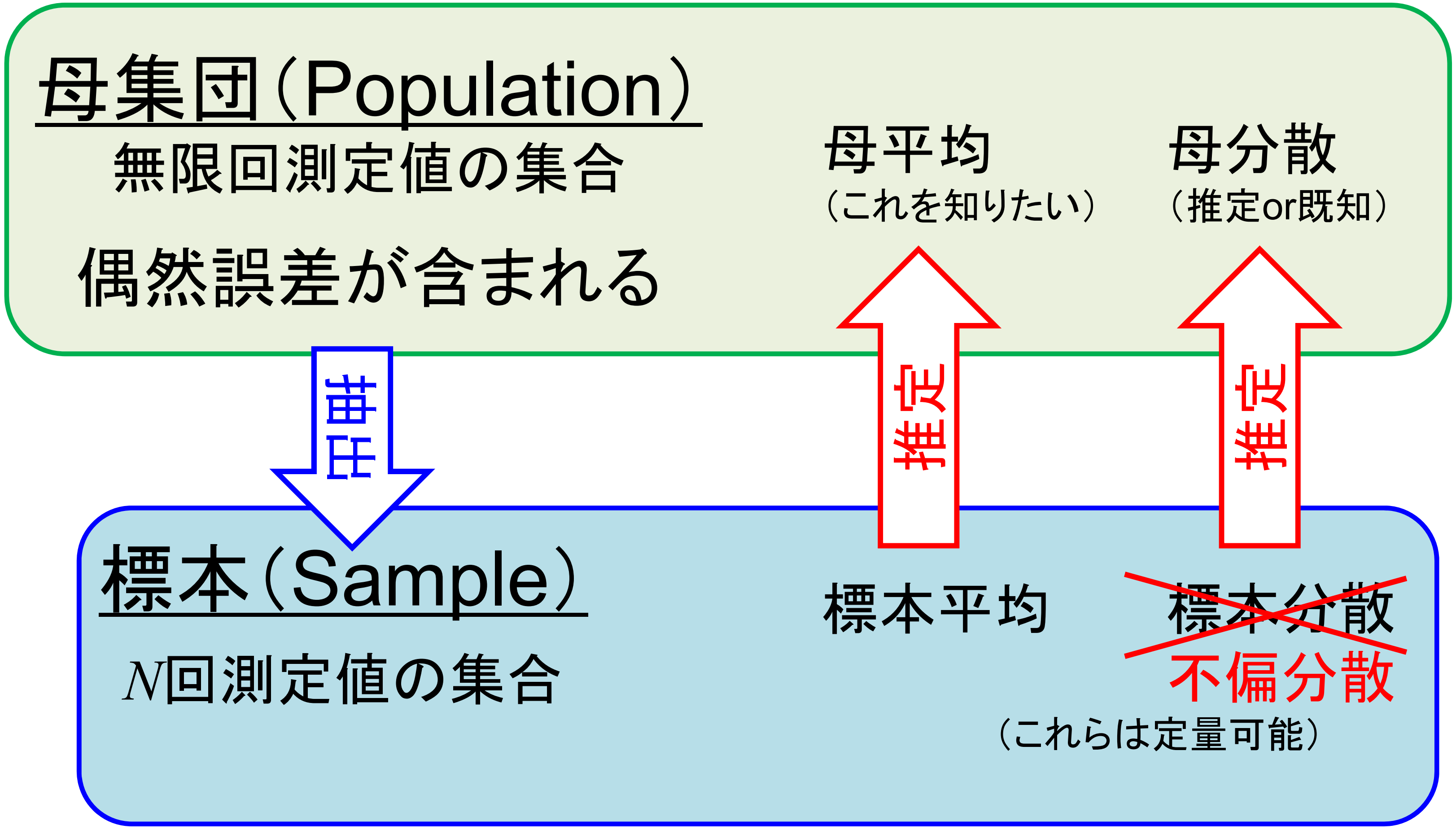
指数関数

$$q \pm \delta q = (A \pm \delta A)\exp(-(B \pm \delta B)t)$$

$$\delta q = \sqrt{(AB\exp(-Bt)\delta B)^2 + (\exp(-Bt)\delta A)^2}$$

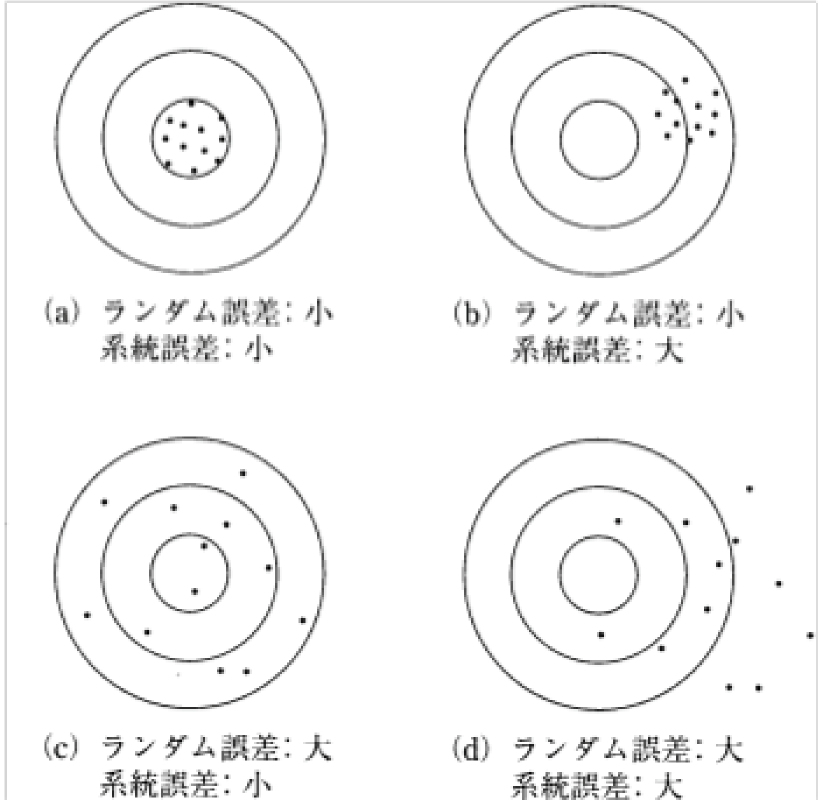
$$= \exp(-Bt)\sqrt{(AB\delta B)^2 + (\delta A)^2}$$

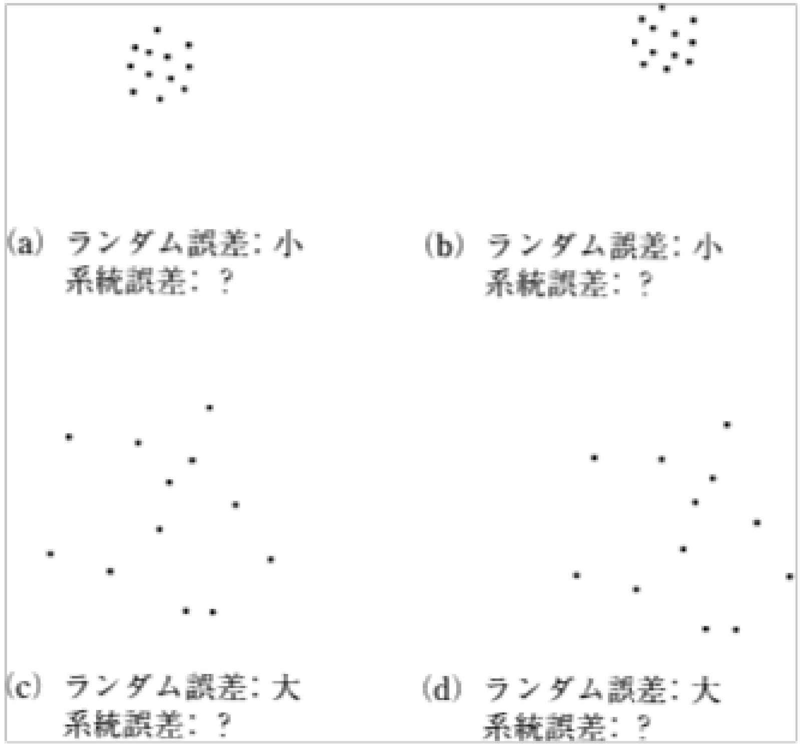
測定の目的



母平均が測定したい値　（系統誤差がなければ）
少ない標本数からどう母平均を推定するか？

何度も測定した結果





真値がわかれば系統誤差、偶然誤差も明確に定義できる。
実際は何度測定しても、ばらつき(偶然誤差)しかわからない

J. R. Taylor (林、馬場　訳) 計測における誤差解析入門、東京化学同人, 2000.
p.101-102図4・1,2より転載。

多数の測定結果　統計処理

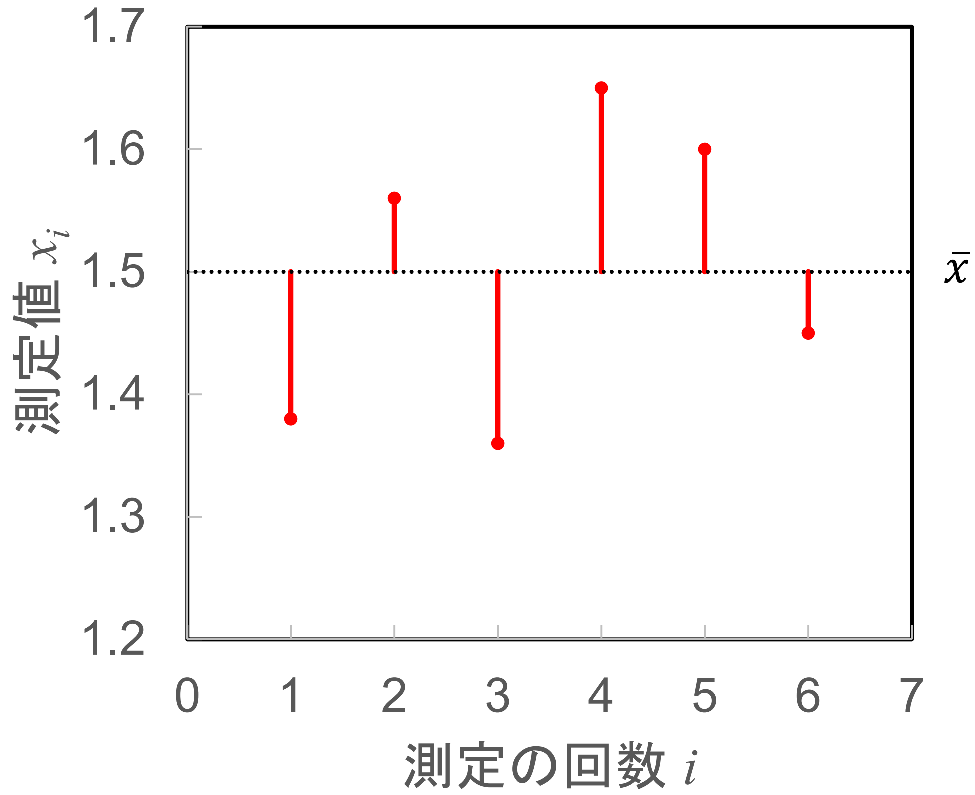
標本平均

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_N}{N} =$$

残差

$$d_i = x_i - \bar{x}$$

i	x_i	d_i
1	1.38	–0.12
2	1.56	0.06
3	1.36	–0.14
4	1.65	0.15
5	1.60	0.10
6	1.45	–0.05
	$\bar{x}$	Σd_i
	1.50	0.00



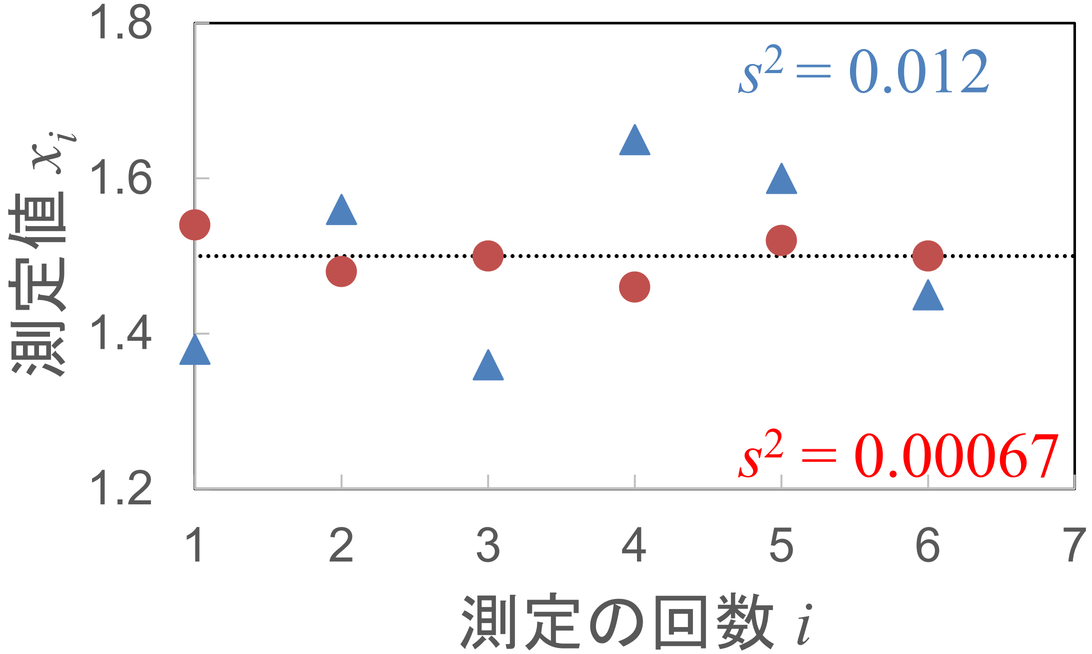
平均するには複数の測定値が必要

どんな標本でも残差の総和は0
⇒ 残差は標本の特徴として不適当

基本的な統計量：標本分散

標本分散

$$s^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2 = \frac{1}{N}\sum_{i=1}^{N} d_i^{\,2}$$



	A	B	C
1	i	x_i	d_i^2
2	1	1.38	0.014
3	2	1.56	0.0036
4	3	1.36	0.020
5	4	1.65	0.022
6	5	1.60	0.010
7	6	1.45	0.0025
8		$\bar{x}$	$s^2 = \Sigma d_i^2/N$
9		1.50	0.012

=VAR.P(C2:C7)

=AVERAGE関数(B2:B7)

値のバラつきを定量化

$d_i^2 \geq 0$なので，
標本分散は常に正の値
データがばらつくほど s^2が大きい

基本的な統計量：標本標準偏差

標本標準偏差　標本分散の平方根

	A	B	C
1	i	x_i	d_i^2
2			
3	1	1.38	0.014
4	2	1.56	0.0036
5	3	1.36	0.020
6	4	1.65	0.022
7	5	1.60	0.010
8	6	1.45	0.0025
9		$\bar{x}$	$s^2 = \Sigma d_i^2/N$
10		1.50	0.012
11			$s = (\Sigma d_i^2/N)^{1/2}$
12			0.11

↑

=STDEV.P(C3:C8)

$$s = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} d_i^{\,2}}$$

データがばらつくほど s も大きい

測定値と同じ単位なので使いやすい

母標準偏差σの不偏推定量ではないことに注意

偶然(ランダム)誤差

多くの原因から生ずる微細な誤差の集合

原因が不明瞭、不規則、補正できない

偶然誤差源の例:

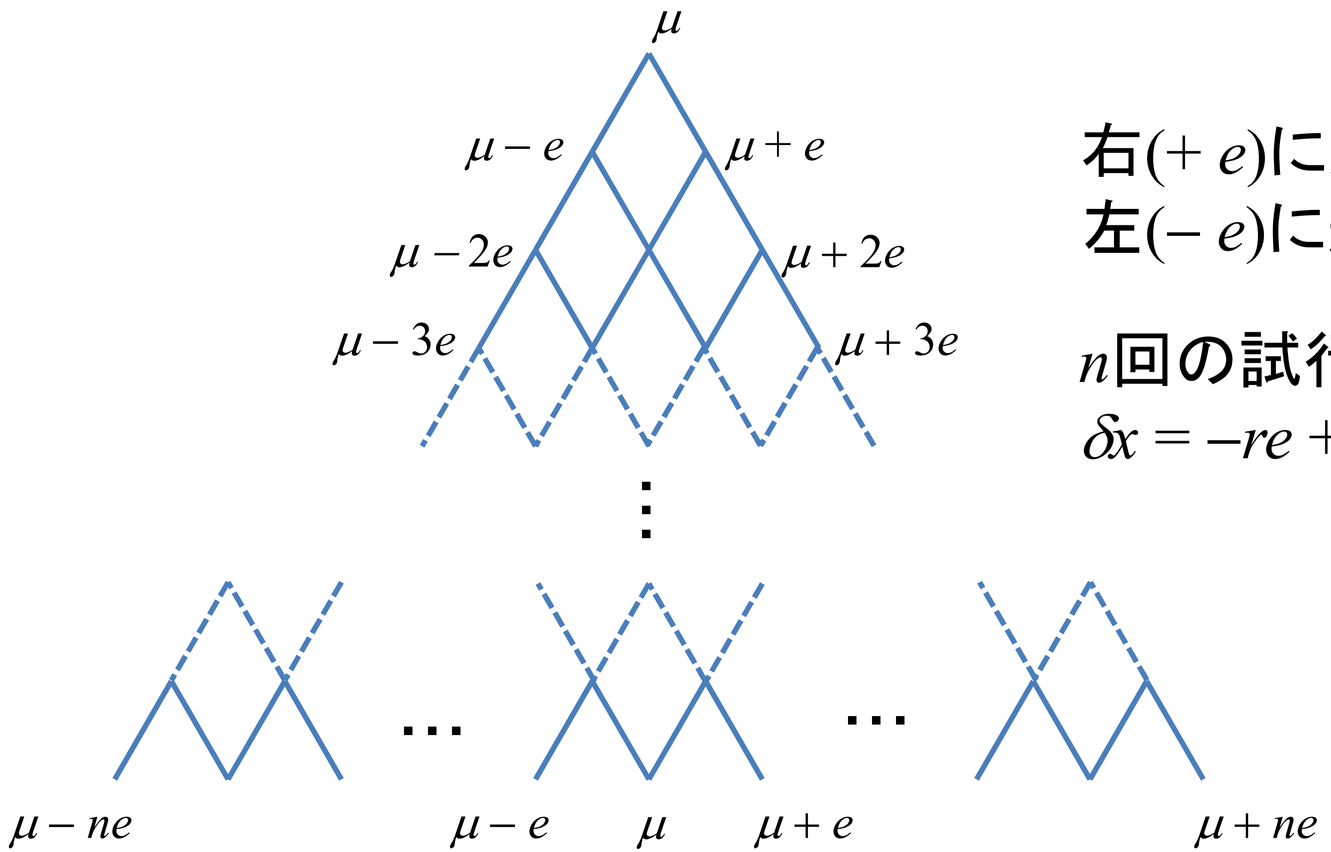
- 気流の乱れによる電子天秤の値の変化
- 熱ノイズ、ショットノイズなど
- 最小目盛の1/10の読み取り誤差
- ストップウォッチを押すタイミング

避けられない(小さくすることは可能)

微細な誤差が数多く重なったとき、
測定値はどのような分布になるか?

偶然誤差の(極めて単純な)モデル

真値がμであるような量xを測定するとき、微小な偶然誤差$\pm e$を与える誤差源がn個ある。

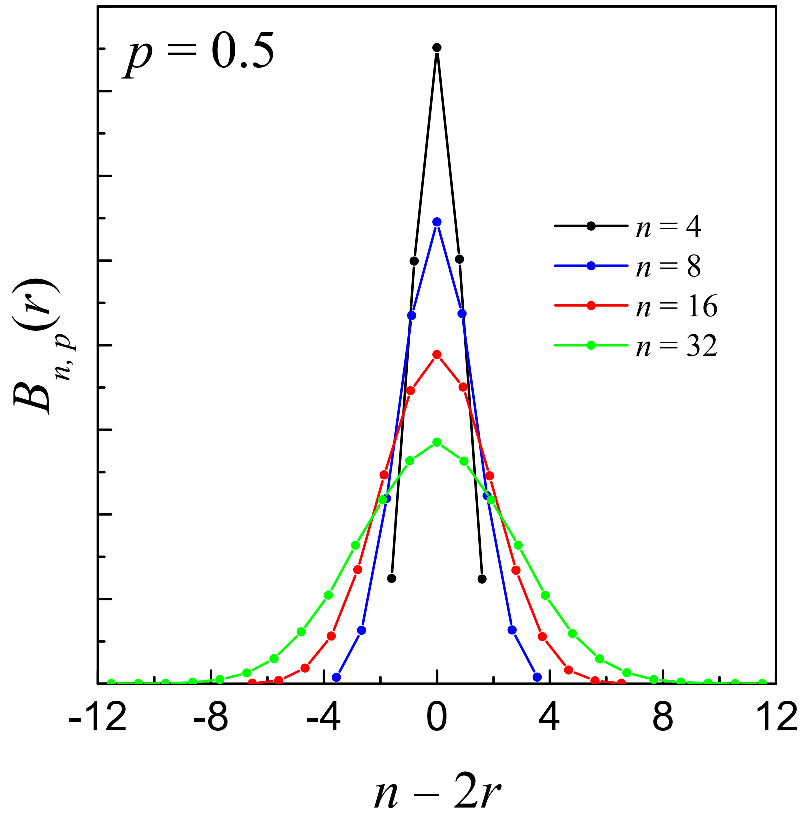


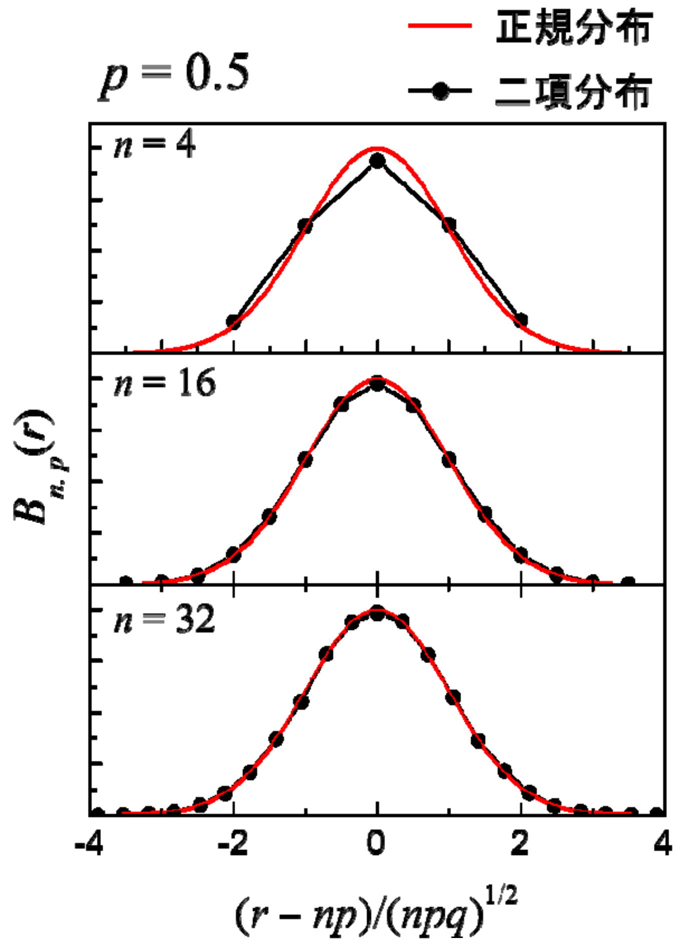
右$(+e)$に進む確率　p
左$(-e)$に進む確率　$q = 1 - p$

n回の試行でr回左に進めば
$\delta x = -re + (n-r)e = (n-2r)e$

n個の誤差を重ね合わせた後の測定値xの分布⇒二項分布

二項分布 $B_{n,p}(r)$から正規分布 $G_{\mu,\sigma}(x)$へ





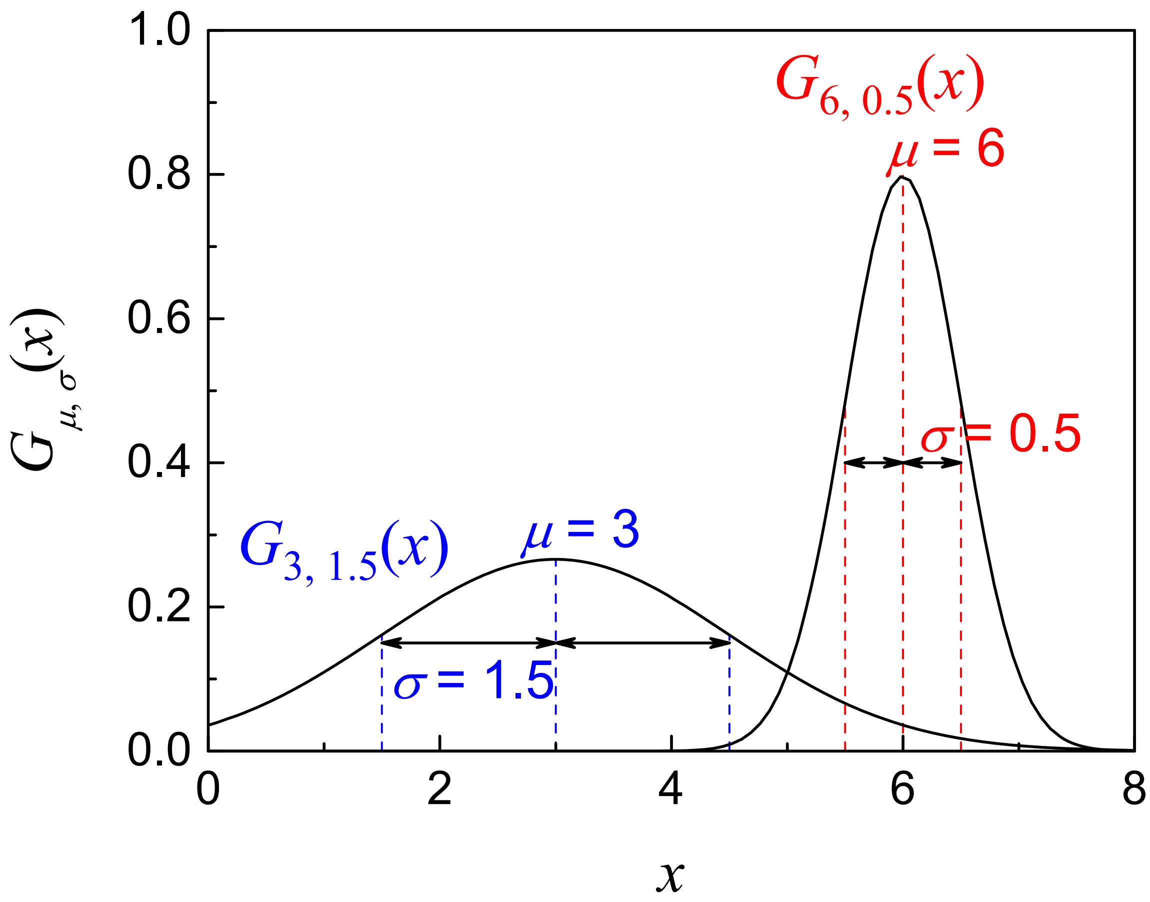
$$\lim_{n\to\infty} B_{n,p}(r) = G_{\mu,\sigma}(x)$$

$$= \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

誤差源の数nを十分大きくすると…正規分布に近づく

正規分布の性質

$$G_{\mu,\sigma}(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$



μ : ピーク位置
　μ = 平均
$\mu \pm \sigma$: 変曲点
　σ = 標準偏差

規格化されている

$$\int_{-\infty}^{\infty} G_{\mu,\sigma}(x)\mathrm{d}x = 1$$

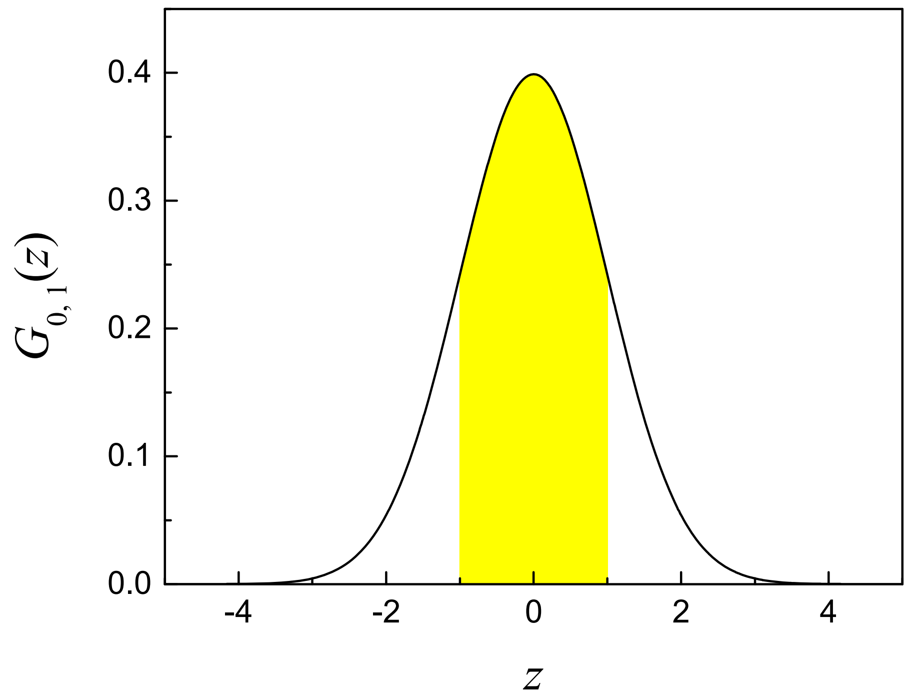
積分値はxが積分範囲内に入る確率を示す

$$\int_{\mu-\sigma}^{\mu+\sigma} G_{\mu,\sigma}(x)\mathrm{d}x$$

$=$ xが$\mu \pm \sigma$に入る確率

標準正規分布 $G_{0,\,1}(x)$

標準化 $\boxed{z = \dfrac{x-\mu}{\sigma}}$ $\quad G_{\mu,\sigma}(x) \Rightarrow G_{0,1}(z) = \dfrac{1}{\sqrt{2\pi}}\exp\left(-\dfrac{z^2}{2}\right)$



$G_{0,1}(z)$: 標準正規分布

$$\int_{\mu-\sigma}^{\mu+\sigma} G_{\mu,\sigma}(x)\mathrm{d}x = \int_{-1}^{1} G_{0,1}(z)\mathrm{d}z$$
$$= 0.68$$

値が標準偏差の範囲(x±σあるいはz±1)に
存在する確率は68%

測定回数が少ない場合

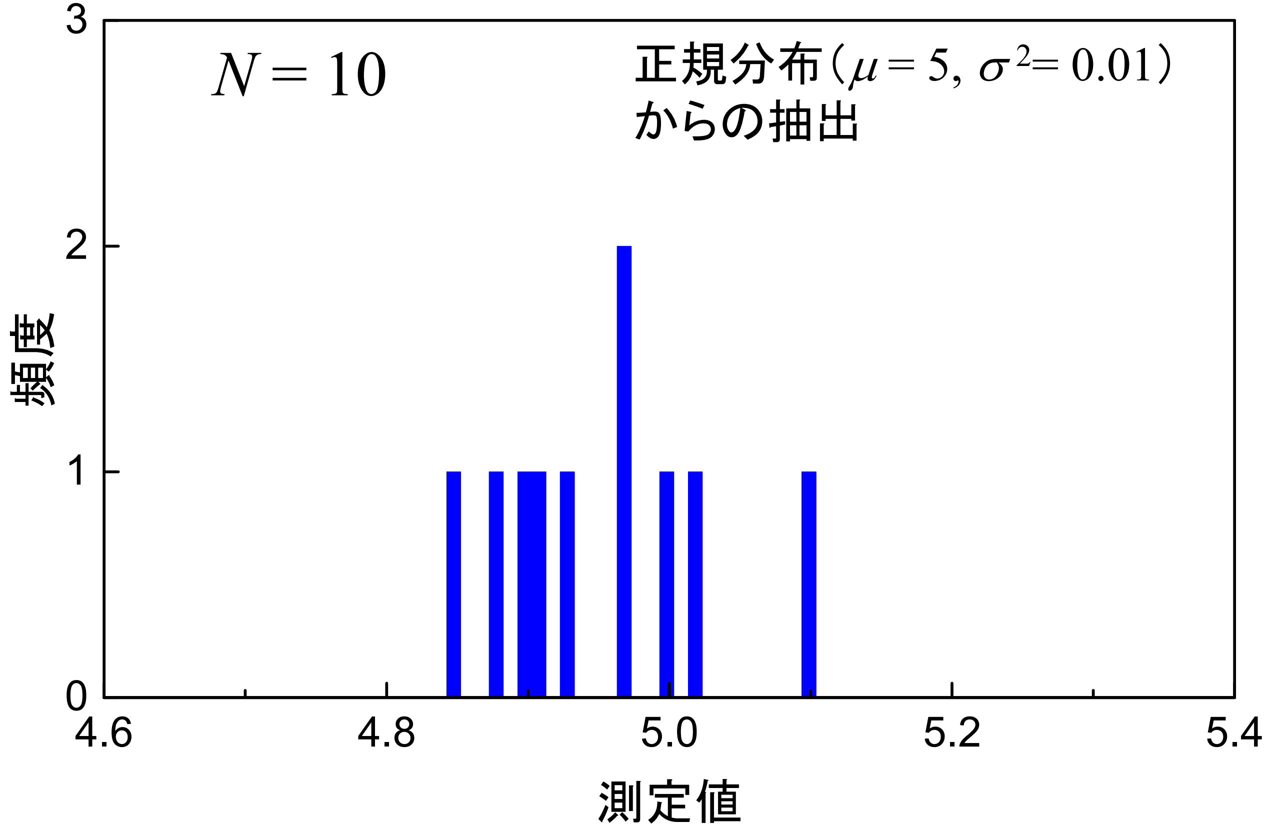
数多くの測定ができれば

母平均μ,母標準偏差σが正確にわかる

測定値
5.02
4.88
4.85
4.97
5.00
4.93
4.91
4.90
4.97
5.10

ヒストグラムにすると…

現実の測定は、せいぜい3–10回



母平均μ（真値）を推定するには？

母平均の推定

標本平均：母平均μ(真値)の最良推定値

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_N}{N} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

正規分布($\mu = 5,\ \sigma^2 = 0.01$)から無作為に抽出

	標本1	標本2	標本3	標本4
1	5.10	4.97	5.05	4.96
2	5.06	5.13	4.86	5.13
3	5.08	4.92	4.92	5.01
4	4.97	4.90	5.00	4.87
5	4.97	5.15	4.97	4.99
標本平均	5.04	5.01	4.96	4.99

標本平均も分布を持つ(値がばらつく)

⇒どのような分布に従うか？

標本平均の分散は母分散ではない

標本平均 $\bar{x} = \dfrac{x_1 + x_2 + \ldots + x_N}{N}$

標本平均の分散 $\sigma_{\bar{x}}{}^2$

$$= \mathrm{E}\left[\left(\bar{x} - \mathrm{E}[\bar{x}]\right)^2\right]$$

$$= \mathrm{E}\left[\left(\frac{x_1 + x_2 + \ldots + x_N}{N} - \mu\right)^2\right]$$

$$= \frac{1}{N^2}\mathrm{E}\left[\left\{(x_1 - \mu) + (x_2 - \mu) \ldots + (x_N - \mu)\right\}^2\right]$$

$$= \frac{1}{N^2}\mathrm{E}\left[\left\{(x_1 - \mu)^2 + (x_2 - \mu)^2 \ldots + (x_N - \mu)^2\right\} + \underline{\left\{(x_1 - \mu)(x_2 - \mu)\ldots\right\}}\right]$$

共分散はゼロ！

$$= \frac{1}{N^2}\left\{N\sigma^2 + 0\right\} = \frac{\sigma^2}{N}$$

$$\boxed{\sigma_{\bar{x}}{}^2 = \frac{\sigma^2}{N}}$$

中心極限定理

平均μ、分散σ^2の”任意分布”に従う母集団
から大きさNの標本を取り出す。

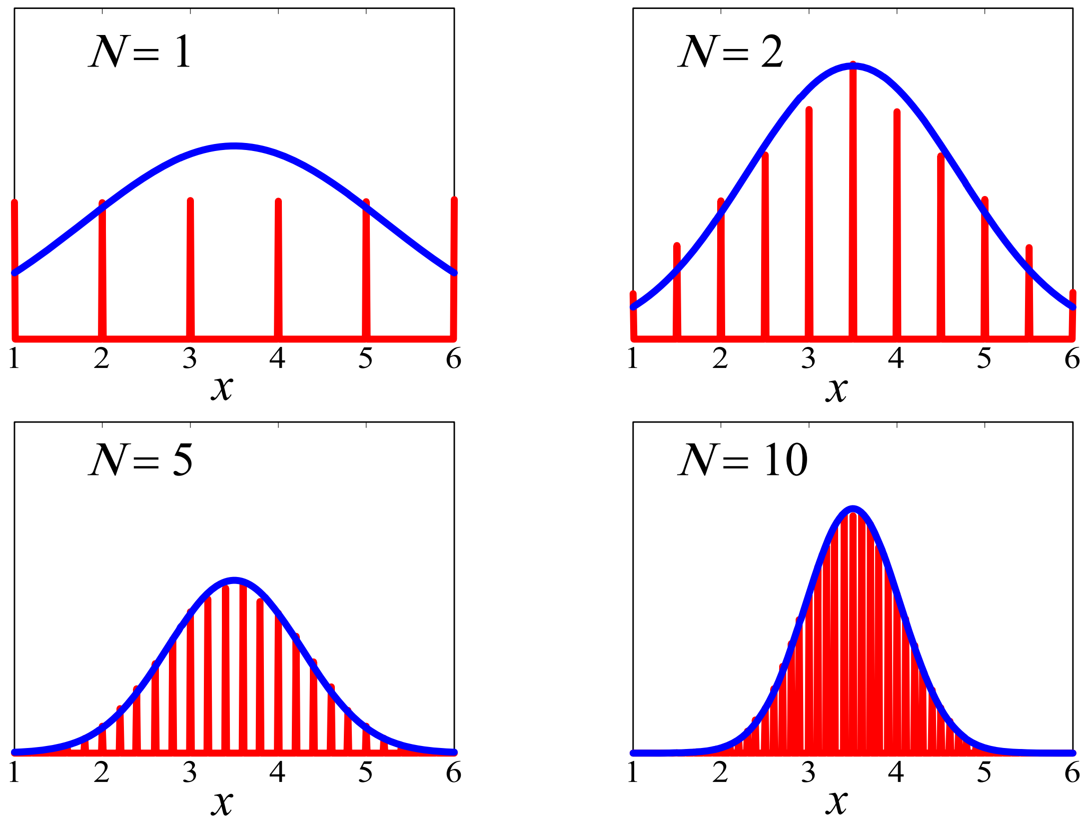
その標本平均はNが十分に大きければ
平均μ、分散σ^2/Nの正規分布　に従う。

観測される現象が多くの誤差源の総和としての誤差を持つ場合、個々の誤差の分布とは無関係に、全体としての誤差は正規分布に従うことが期待される。

「偶然誤差は正規分布に従う」

Nが十分に大きくない場合でもよい近似
目安　　母集団が左右対称な分布なら $N > 4$ 程度
非対称分布なら$N > 30$ 程度

サイコロにおける中心極限定理



同時に振るサイコロの数を増やすと，標本平均の分布は正規分布に近づく

赤線：複数のサイコロを振ったときの標本平均の分布
試行回数100000回
青線：$G_{3.5,\sqrt{35/12N}}$で表現される正規分布

標本平均の分布

N回測定の平均値 $\bar{x}$ の分布

正規分布 $G_{\mu,\sigma/\sqrt{N}}(\bar{x})$

- 標本平均の平均

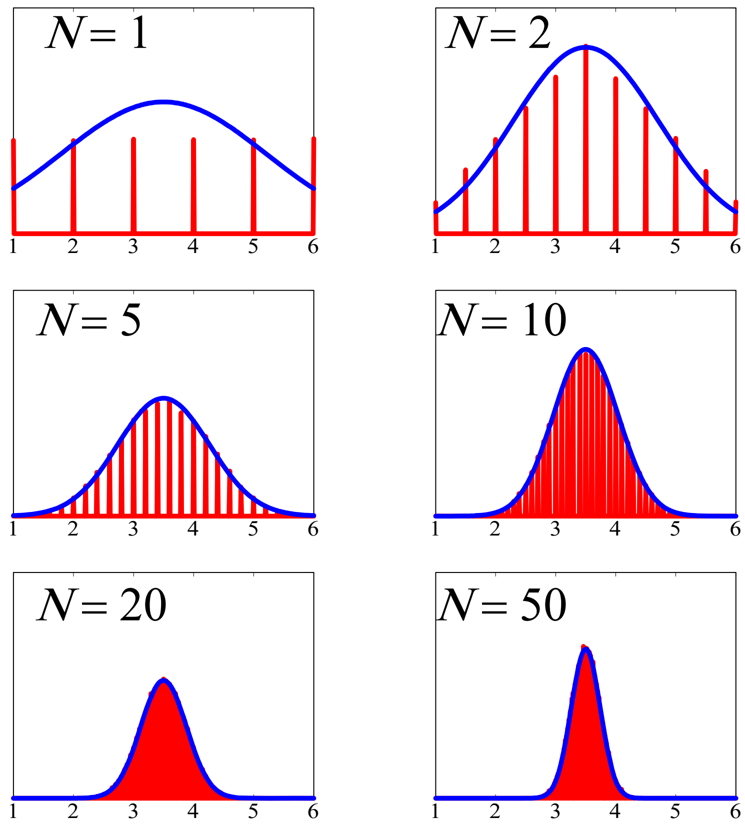
$$\boxed{\mu_{\bar{x}} = \mu}$$

母平均と一致

- 標本平均の分散

$$\boxed{\sigma_{\bar{x}}{}^2 = \frac{\sigma^2}{N}}$$

標本サイズに反比例して小さくなる



赤線:サイコロの目の平均値の分布
青線: $G_{3.5,\sqrt{35/12N}}(x)$

サイコロの目は$\mu = 3.5$, $\sigma^2 = 35/12$
Nが大きい場合の平均値の分布は
$G_{3.5,\sqrt{35/12N}}(x)$と一致する

母分散σ^2の最良推定値は？

σ^2が未知の場合、標本から推測しなければならない

標本分散s^2はσ^2の最良推定値？

$$s^2 = \frac{1}{N}\sum(x_i - \bar{x})^2$$

正規分布（$\mu = 5,\ \sigma^2 = 0.01$）から無作為に抽出

	標本1	標本2	標本3	標本4
1	5.10	4.97	5.05	4.96
2	5.06	5.13	4.86	5.13
3	5.08	4.92	4.92	5.01
4	4.97	4.90	5.00	4.87
5	4.97	5.15	4.97	4.99
s^2	0.0031	0.011	0.0043	0.0071
s	0.055	0.11	0.065	0.084

s^2ではσ^2を過小評価？

標本分散は偏っている

- 偏差の二乗の平均値

$$\boxed{\sigma^2 = \frac{1}{N}\sum(x_i - \mu)^2} \quad (1)$$

<u>母平均μとの差（偏差）</u>

母平均と標本平均の違い

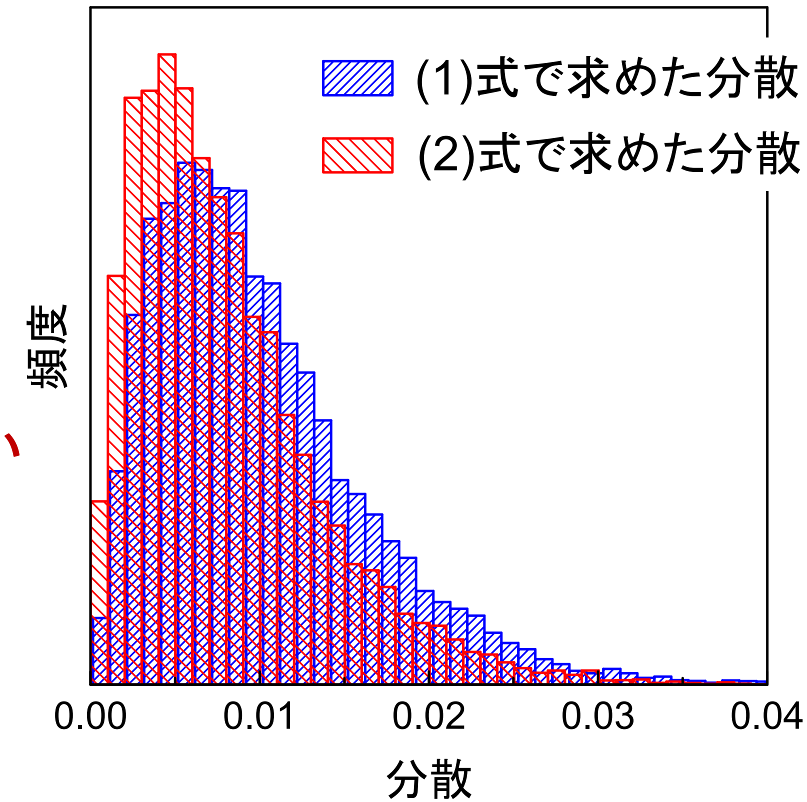
- 残差の二乗の平均値

$$\boxed{s^2 = \frac{1}{N}\sum(x_i - \bar{x})^2} \quad (2)$$

<u>標本平均 $\bar{x}$ からの残差</u>

$\bar{x}$ は$\Sigma(x_i - \bar{x})^2$を最小にするような値

$\Rightarrow s^2$はσ^2より小さめに偏る



正規分布（$\mu = 5,\ \sigma^2 = 0.01$）から
$N = 5$の標本を繰り返し取得

(1)式（母分散）　平均：0.01
(2)式（標本分散）平均：0.008

不偏分散：母分散の最良推定値

不偏分散

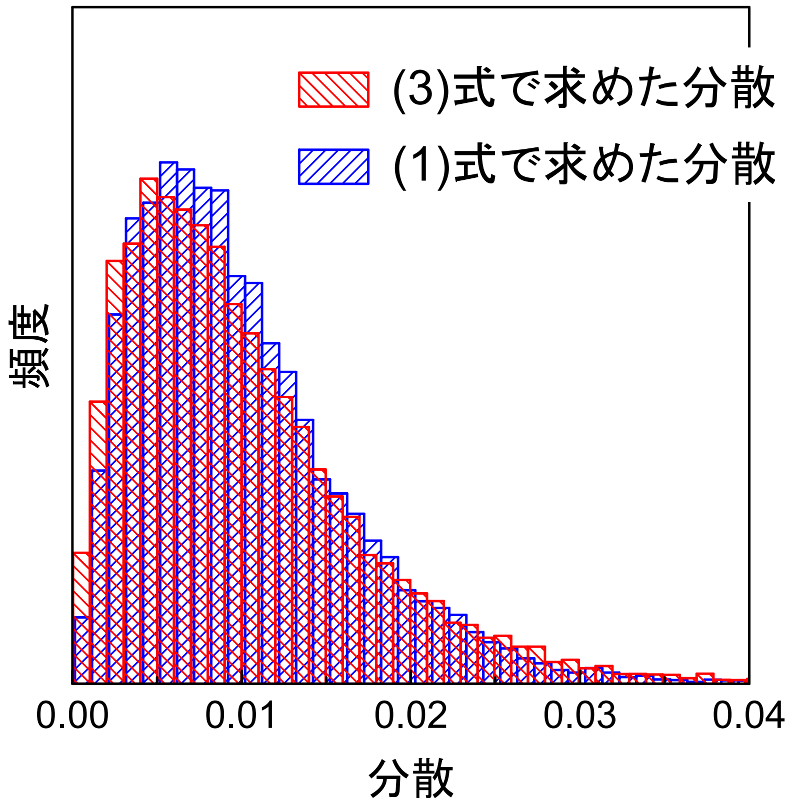
Nの代わりに$N-1$で割る

$$\boxed{u^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2} \quad (3)$$

$N-1$を**自由度**という

標準偏差

$$u = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2}$$



正規分布（$\mu = 5$, $\sigma^2 = 0.01$）から
$N = 5$の標本を繰り返し抽出

(1)式（母分散）　平均：0.01
(3)式（不偏分散）平均：0.01

不偏分散を用いた標準化

標本平均 $\bar{x}$ ⇐正規分布

<u>σ^2が既知</u>

⇨ $z = \dfrac{\bar{x} - \mu}{\sigma / \sqrt{N}}$ ⇐標準化すると標準正規分布に

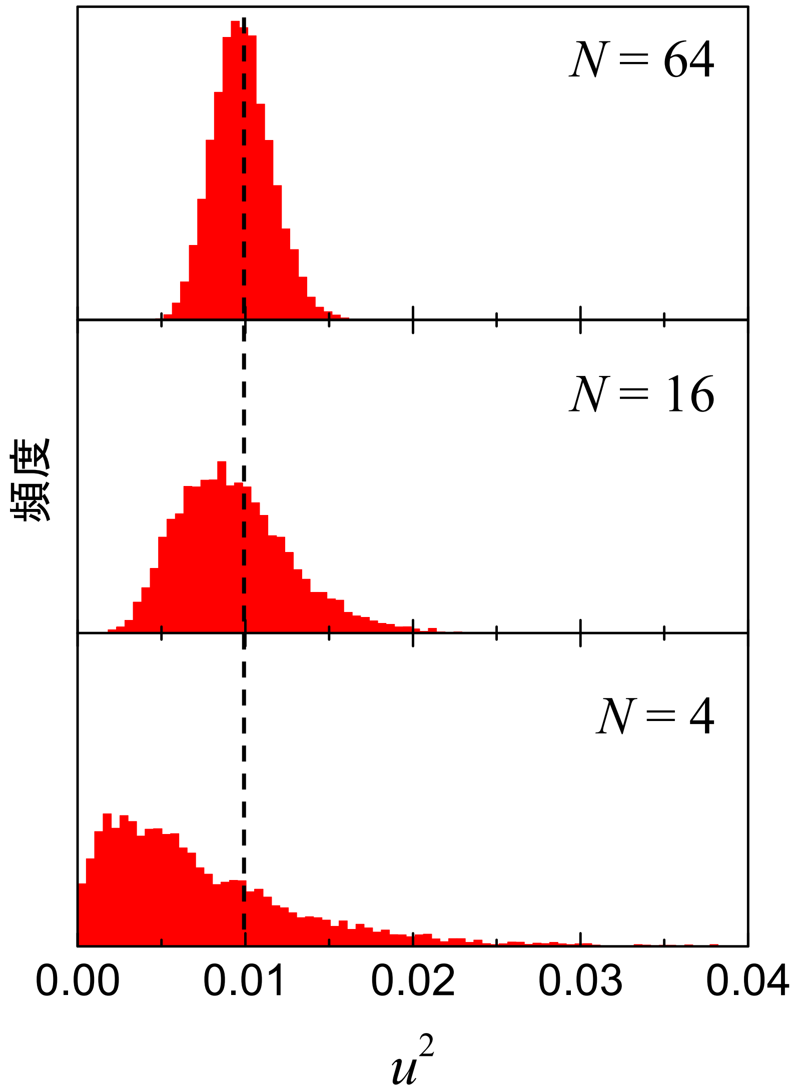
<u>σ^2が未知…最良推定値u^2を使う</u>

⇨ $t = \dfrac{\bar{x} - \mu}{u / \sqrt{N}}$ ⇐標準化しても標準正規分布に従わない

σ^2は定数、<u>u^2は分布を持つ</u>

不偏分散の分布

正規分布（$\mu = 5,\ \sigma^2 = 0.01$）から
N個の標本を繰り返し抽出



標本数が少ない場合
u^2は平均値に対して**非対称**な分布

u^2がその平均値より低い確率

$N = 4$	59 %
$N = 16$	55 %
$N = 64$	52 %

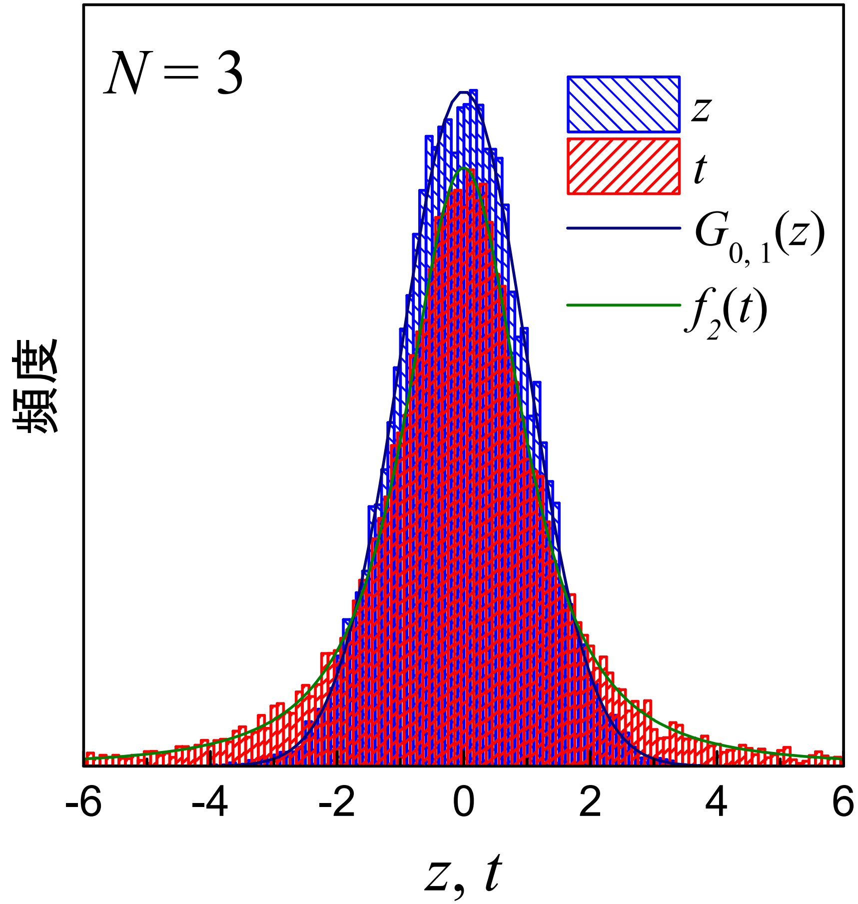
⇨ σ^2を過小評価しがち

$$\boxed{t = \frac{\bar{x} - \mu}{u / \sqrt{N}}}$$

標準正規分布よりも広がった分布

自由度N–1のt-分布

Studentのt-分布



正規分布（$\mu = 5$, $\sigma^2 = 0.01$）から
$N = 3$の標本を繰り返し抽出

- σ^2を用いて標準化(z)
- u^2を各標本で計算・標準化(t)

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{N}}$$

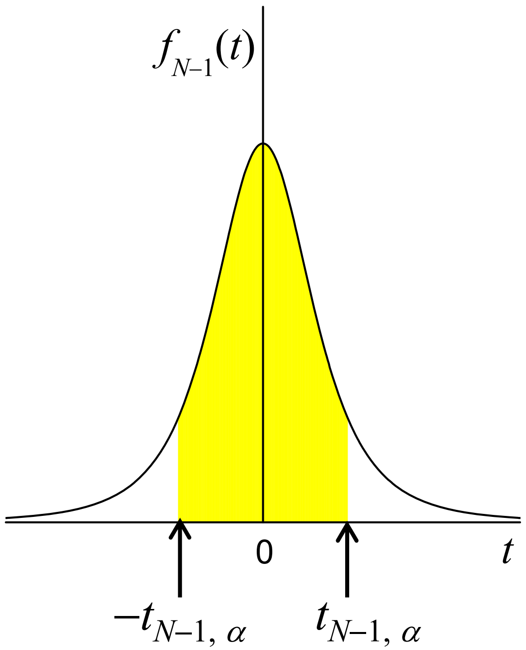
$$t = \frac{\bar{x} - \mu}{u / \sqrt{N}}$$

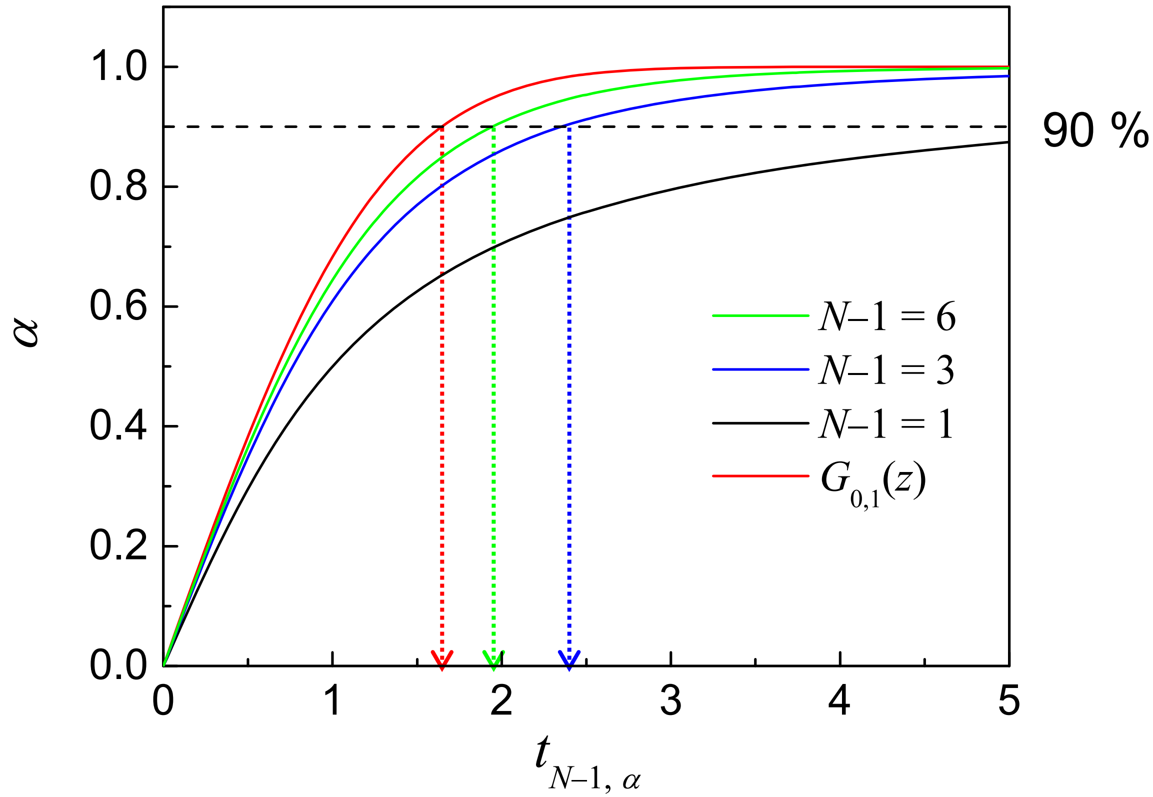
t-分布は標準正規分布に比べ幅が広い
（特に自由度が小さいとき）

t-分布は自由度が大きくなると、標準正規分布に近づく

信頼区間

$t = \dfrac{\bar{x} - \mu}{\sqrt{u^2 / N}}$ が$\pm\, t_{N\text{-}1,\alpha}$の間にある確率…　　$\alpha = \displaystyle\int_{-t_{N-1,\alpha}}^{t_{N-1,\alpha}} f_{N-1}(t)\mathrm{d}t$





同じαでも自由度が小さいほど信頼区間が広い

t-検定表

表6.3　Student の t 分布における $t_{N-1}(z\ \%)$ の値

表中の"P（有意水準），両側検定"については，8.2で述べる．

自由度 $(N-1)$	標本数 (N)	$t_{N-1}(z\ \%)$, $z(\%)$：信頼度			
		68.3%	90%	95%	99%
		P(有意水準)，両側検定			
		31.7%	10%	5%	1%
1	2	1.837	6.314	12.706	63.657
2	3	1.321	2.920	4.303	9.925
3	4	1.197	2.353	3.182	5.841
4	5	1.141	2.132	2.776	4.604
5	6	1.110	2.015	2.571	4.032
6	7	1.090	1.943	2.447	3.707
7	8	1.077	1.895	2.365	3.500

Excelでは、T.INV.2T関数

信頼区間の決め方

標本平均 $\bar{x}$ ⇐正規分布

不偏分散$u^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2$ ⇐ 分布を持つ

⇨ $t = \dfrac{\bar{x} - \mu}{u / \sqrt{N}}$ ⇐自由度N–1のt分布

⇨ tは α %の確率で $-t_{N-1,\alpha} \leq t \leq t_{N-1,\alpha}$

⇨ μは α %の確率で $\bar{x} - t_{N-1,\alpha}\dfrac{u}{\sqrt{N}} \leq \mu \leq \bar{x} + t_{N-1,\alpha}\dfrac{u}{\sqrt{N}}$

$\pm t_{N-1,\alpha}\dfrac{u}{\sqrt{N}}$:信頼区間(エラーバー)、α %:信頼度

誤差の表記

母標準偏差σが**未知**

$$x = \bar{x} \pm t_{N-1,\alpha}\frac{u}{\sqrt{N}}$$

標準誤差（α=0.68のとき）

$$x = \bar{x} \pm t_{N-1,0.68}\frac{u}{\sqrt{N}}$$

母平均は68%の確率でこの範囲内に入る

27th May, 2017

検出限界と定量下限、信頼性に関わる用語

明星大学大学院理工学研究科　上本道久

1. それぞれの用語の定義と使い分け
2. 検出限界の求め方
3. 定量下限の求め方
4. 信頼性に関わる用語

– 分野により異なる用語の実態

1．検出限界、定量下限および感度の定義

- 検出限界（検出下限）(limit of detection, LOD, limit of identification) とは検出できる最小量（値）
- 定量下限（定量限界とはいわない）(minimum limit of determination)とは、ある分析方法で目的物質の定量が可能な最小量（値）又は濃度

検出限界 « 定量下限

- 感度(sensitivity)とは、

ⓐ検出下限で表した分析方法あるいは機器の性能

ⓑ検量線の傾きで表した分析方法の性能（ある量を検出するときに検出定量できる被測定量の変化の最小量（値））

JIS K 0211:2013「分析化学用語（基礎部門）」

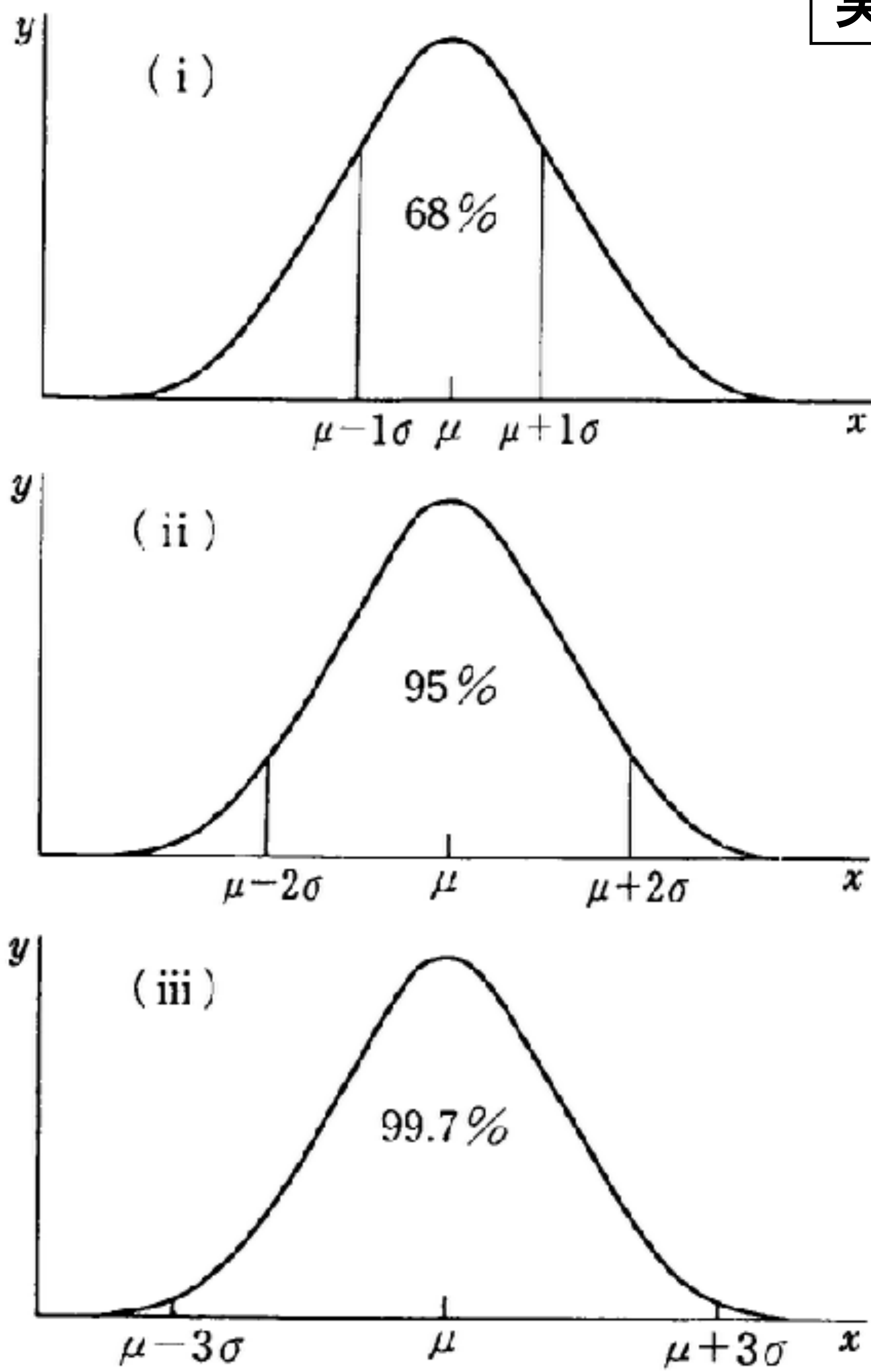
- 検出限界：検出可能な最小量

- 検出可能な最小量；
ブランク（測定対象成分を含まない試料）と有意に区別できる信号を与える最小量

- ブランクと有意に区別できる信号；
ブランクの平均値とばらつき度合（分布）から決定

検出限界値を求めるためには、ブランク試料での平均値とその分布に関する特性値（標準偏差など）が必要

実測値は分布（ばらつき）を有する



正規分布の特性

母集団の
(i)約68%は平均±1σ以内にある

(ii)約95%は平均±2σ以内にある

(iii)約99.7%は平均±3σ以内にある

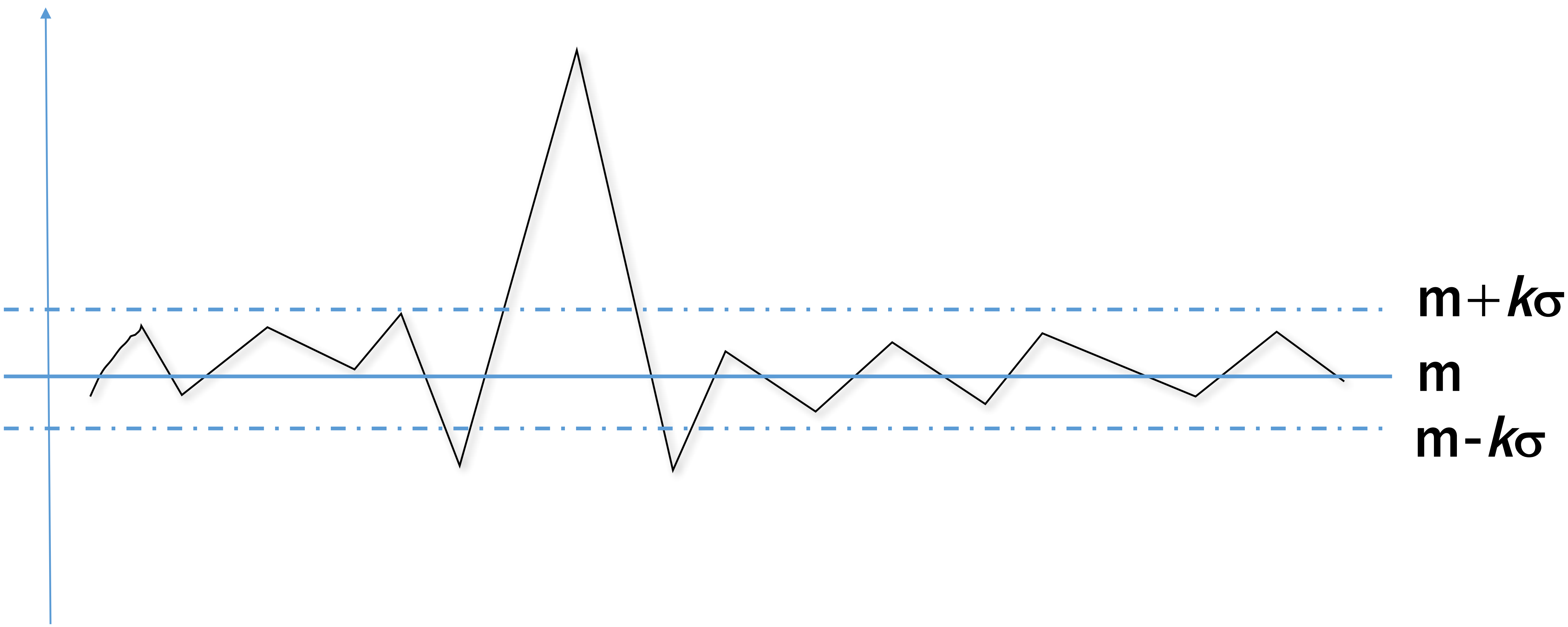
標準偏差；nが大きくない（例えばn<10）とき、
不偏平均二乗偏差（不偏分散）の平方根とするのが一般的

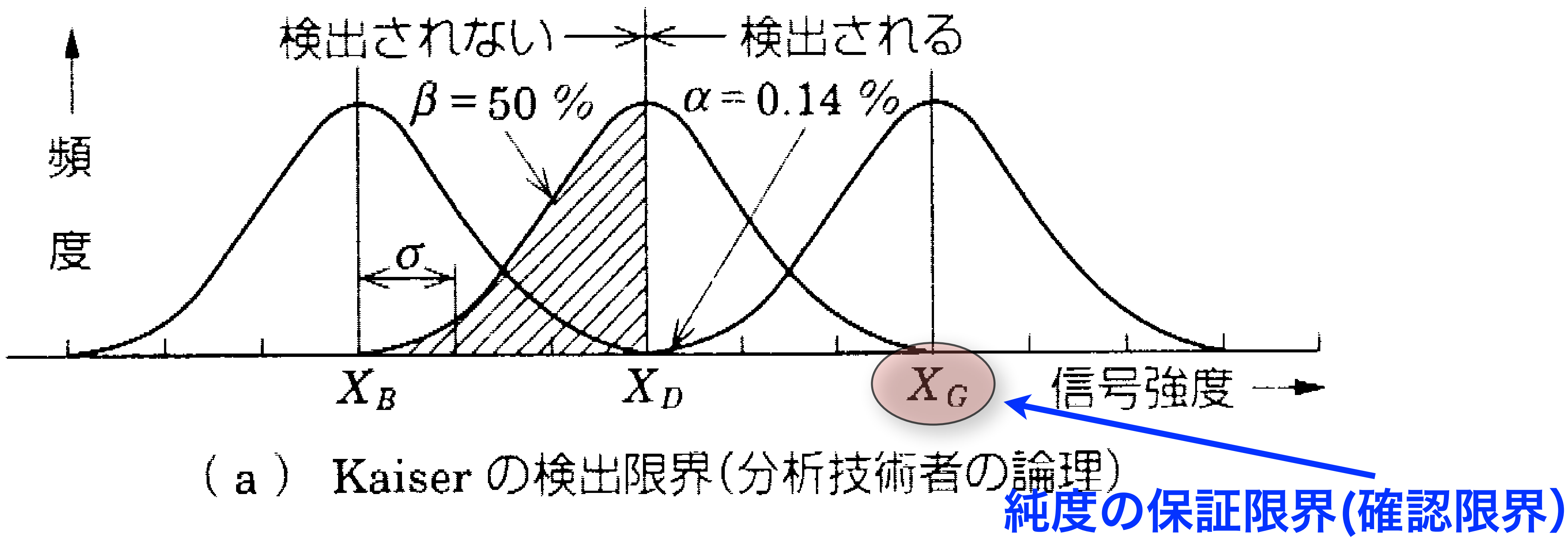
$$\sigma = \sqrt{\frac{\Sigma\left(x_i - \bar{x}\right)^2}{n-1}}$$

平均偏差；nが大きくない（例えばn<10）ときは、
偏差の絶対値の相加平均も用いられる

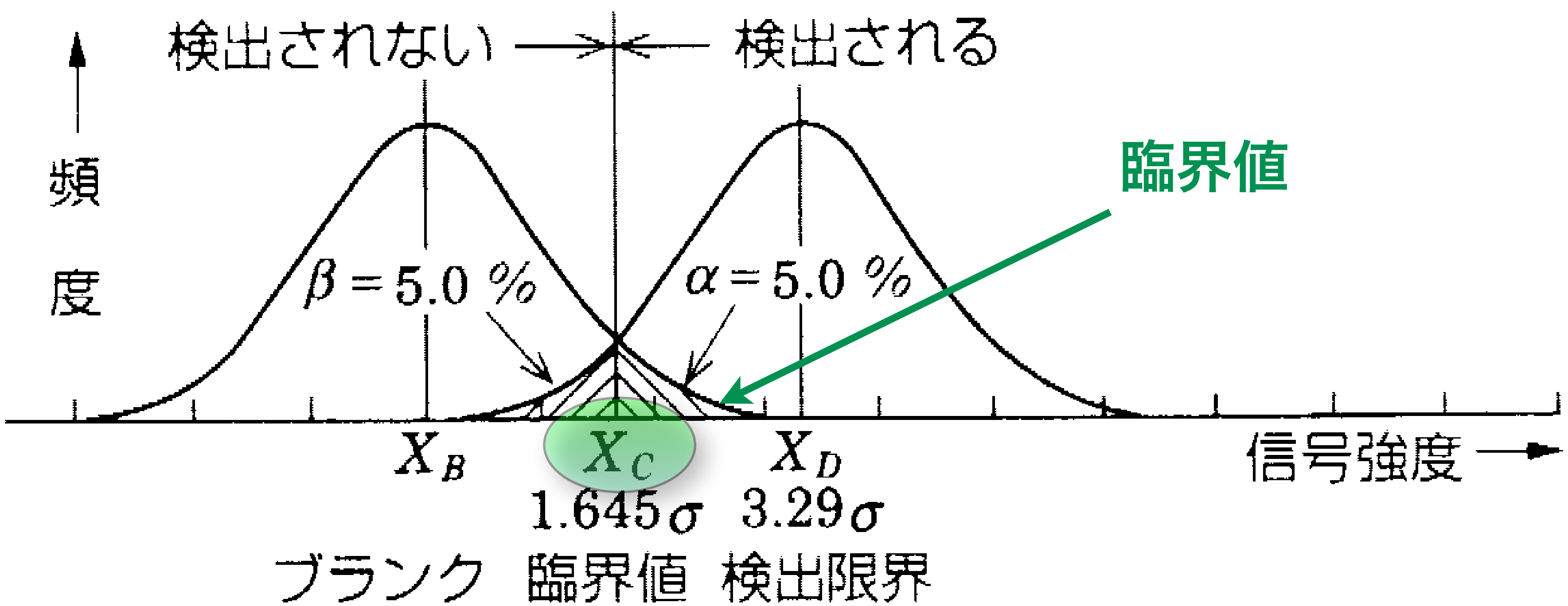
$$\tau = \frac{\Sigma\,|\,x_i - \bar{x}\,|}{n}$$

2. 検出限界の求め方





（a） Kaiser の検出限界(分析技術者の論理)



（b） Currie の検出限界(分析管理者の論理)

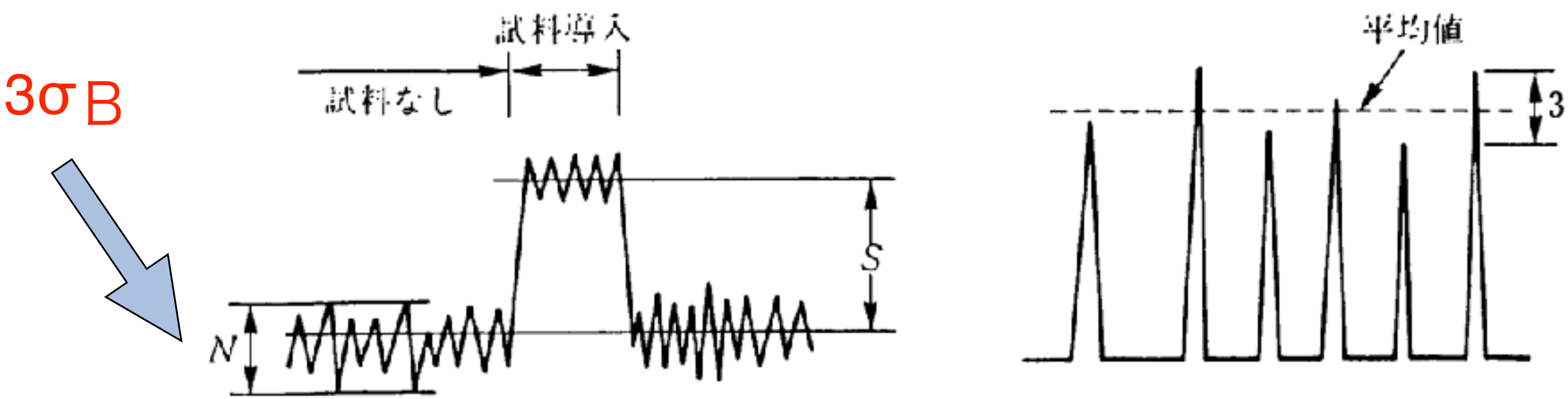
X_B; ブランク値の平均、 X_D; 検出限界

検出限界に関わる用語

- 検出限界(Limit of detection)$_1$　3.29σ
  $\alpha = \beta = 0.05$
- 検出限界(Limit of detection)$_2$　3σ
  $\alpha = 0.0014$, $\beta = 0.50$ ($\alpha = \beta = 0.067$)
- 確認限界(Limit of decision)　1.645σ
  $\alpha = 0.05$, $\beta = 0.50$ ($\alpha = \beta = 0.206$)
- 純度の保証限界(確認限界)
  (Limit of guarantee of purity)　6σ
  $\alpha = \beta = 0.0014$

$$(\text{検出限界値}) = 3 \times \sigma_B \times \frac{C}{\overline{X}_S - \overline{X}_B}$$

X_S；測定元素の発光強度の平均値　X_B；バックグラウンドの発光強度の平均値
C ；測定元素の濃度　σ_B；バックグラウンドの標準偏差で、10回程度測定を繰返して統計的に求める。



(a) 定常的信号の場合　　(b) 過渡的信号の場合

ICP発光分析における分析線と検出限界（一部）

b.　ICP 発光分光分析で用いられる分析線の波長と検出限界

元素	波長[a] nm		軸方向観測検出限界[b] ng mL^{-1}	横方向観測検出限界[b] ng mL^{-1}	元素	波長[a] nm		軸方向観測検出限界[b] ng mL^{-1}	横方向観測検出限界[b] ng mL^{-1}
Ag	I	328.068	0.6	1	Ca	II	396.847	0.01	
Al	II	167.081	0.5	3	Cd	II	214.438	0.3	0.6
Al	I	308.216		8	Cd	II	226.502	0.2	0.6
Al	I	396.153	1.5	3	Cd	I	228.802	0.4	0.5
As	I	189.042	2		Ce	II	413.747	2.4	5
As	I	193.759	5	5	Ce	II	418.660	2	3
Au	I	242.795	1.4	3	Cl	I	134.724	19	
B	I	182.641	0.5		Cl	I	135.166	50	
B	I	249.773	0.5	0.3	Cl	I	837.597		50 ng[c]
Ba	II	455.404	0.05	0.09	Co	II	228.616	0.3	0.8
Be	I	234.861	0.05	0.08	Cr	II	205.552	0.3	1
Be	II	313.042	0.09	0.1	Cr	II	267.716	0.5	0.8
Bi	II	153.317	8.4		Cs	I	455.536	1500	10 000
Bi	I	223.061	5.9	5	Cu	I	324.754	1.3	1
Br	I	148.845	34		Cu	I	327.396	0.6	
Br	I	154.065	9		Dy	II	353.171	0.3	2
Br	I	700.521	3500		Er	II	323.059	0.9	
Br	I	827.246		50 ng[c]	Er	II	337.275	0.4	2
C	I	193.090		10	Er	II	349.910	0.4	
C	I	247.857		40	Eu	II	381.966	0.1	0.09
Ca	II	183.801	1.2		Eu	II	412.974	0.1	
Ca	II	393.367	0.04	0.1	Eu	II	420.505	0.06	

日本分析化学会編：分析化学データブック、改訂5版、2004（丸善）

a.　原子吸光分析の検出限界

元　素	波　長 nm	FAAS[a)]		GFAAS[b)]
		フレーム[c)]	検出限界 μg mL^{-1}	検出限界 pg
Ag	328.1	A/A	0.003	0.3
Al	309.3	N/A	0.03	1
As	193.7	A/A, Ar/H	0.05 (0.1)[d)]	8
Au	242.8	A/A	0.02	2
B	249.8	N/A	2	200
Ba	553.6	N/A	0.03	6
Be	234.9	N/A	0.002	0.03
Bi	223.1	A/A	0.005 (0.2)[d)]	5
Ca	422.7	A/A, N/A	0.002	0.4
Cd	228.8	A/A	0.002	0.08
Co	240.7	A/A	0.008	2
Cr	357.9	A/A	0.005	2
Cs	852.1	A/A	0.05	6
Cu	324.8	A/A	0.005	1
Dy	404.6	N/A	0.2	
Dy	421.2	N/A	0.05	40
Er	400.8	N/A	0.05	80
Eu	459.4	N/A	0.02	20
Fe	248.3	A/A	0.004	4
Ga	294.4	N/A	0.4	20

原子吸光分析における分析線と検出限界（一部）

日本分析化学会編：分析化学データブック、改訂5版、2004（丸善）

元素	波長 (nm)	感度 (μg ml⁻¹1%)	検出限界 (μg ml⁻¹)	元素	波長 (nm)	感度 (μg ml⁻¹1%)	検出限界 (μg ml⁻¹)
Ag	328.1	0.08	0.005	Mo	313.2	1	0.1
Al	309.3	1.1	0.1	Na	589.0	0.04	0.005
As	193.7	1.0	0.2	Nb	334.4	21	5
Au	242.8	0.5	0.02	Nd	463.4	10	2
B	249.8	35	6	Ni	232.0	0.1	0.005
Ba	553.6	0.4	0.05	Os	290.9	1	1
Be	234.9	0.03	0.002	Pb	283.3	0.5	0.01
Bi	223.1	0.7	0.05	Pd	247.6	0.3	0.02
Ca	422.7	0.03	0.002	Pr	495.1	13	10
Cd	228.8	0.03	0.005	Pt	265.9	2	0.1
Co	240.7	0.1	0.005	Rb	780.0	0.2	0.005
Cr	357.9	0.15	0.005	Re	346.0	20	1.5
Cs	852.1	0.5	0.05	Rh	343.5	0.35	0.03
Cu	324.7	0.1	0.005	Ru	349.9	2	0.3
Dy	421.2	0.7	0.4	Sb	217.5	1	0.2
Er	400.8	0.85	0.1	Sc	391.2	0.5	0.2
Eu	459.4	0.75	0.2	Se	196.0	2	0.5
Fe	248.3	0.15	0.005	Si	251.6	1.2	0.1
Ga	287.4	2.3	0.1	Sm	429.7	8.5	5
Gd	368.4	17	4	Sn	224.6	1.2	0.06
Ge	265.2	2.5	1	Sr	460.7	0.2	0.01
Hf	307.3	15	15	Ta	271.5	30	6
Hg	253.7	15	0.5	Tb	432.6	7.5	2
Ho	410.4	1.4	0.3	Te	214.3	2	0.3
In	303.9	0.9	0.05	Ti	364.3	1.4	0.2
Ir	264.0	13	2	Tl	276.8	0.2	0.8
K	766.5	0.1	0.005	V	318.4	1.3	0.04
La	550.1	30	2	W	400.9	35	3
Li	670.8	0.07	0.005	Y	407.7	1.1	0.3
Lu	331.2	15	3	Yb	398.8	0.17	0.04
Mg	285.2	0.008	0.0005	Zn	213.6	0.04	0.002
Mn	279.5	0.08	0.003	Zr	360.1	20	5

フレーム原子吸光分析における感度と検出限界(2)

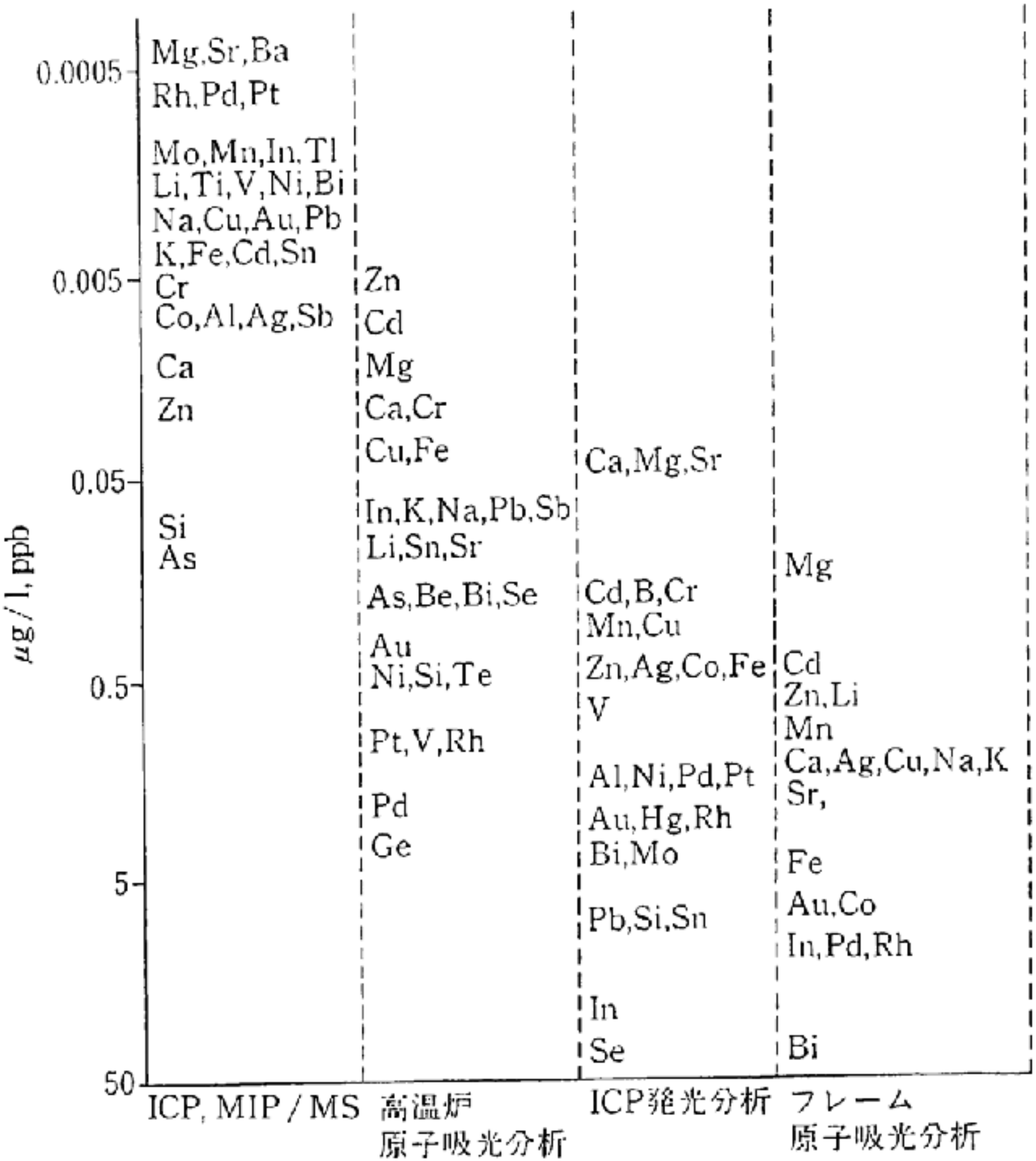
（感度の値が記載されている）

ICP質量分析における測定質量数と検出限界（一部）

		四重極質量分析計			高分解能質量分析計		
元素	質量	Normal[a] 検出限界 / pg mL^{-1}	Cool[b] 検出限界 / pg mL^{-1}	Collision[c] 検出限界 / pg mL^{-1}	分解能[d]	Normal[e] 検出限界 / pg mL^{-1}	Cool[f] 検出限界 / pg mL^{-1}
Ag	107	0.7		0.2	300		0.1
Al	27	0.2	0.08	4	300		0.1
As	75	3		2			
Au	197	0.5		0.4	300		0.08
B	11	0.8		17	300	0.2	
Ba	138	0.3			300	0.02	
Be	9	0.9		0.6			
Bi	209	0.2		0.03	300	0.01	
Br	79	16.5		70			
C	12	50 000		40 000			

日本分析化学会編：分析化学データブック、改訂5版、2004（丸善）

各種分析法による
検出限界の比較



3. 定量下限の求め方

必要とされる精度（ここではRSD）に依存

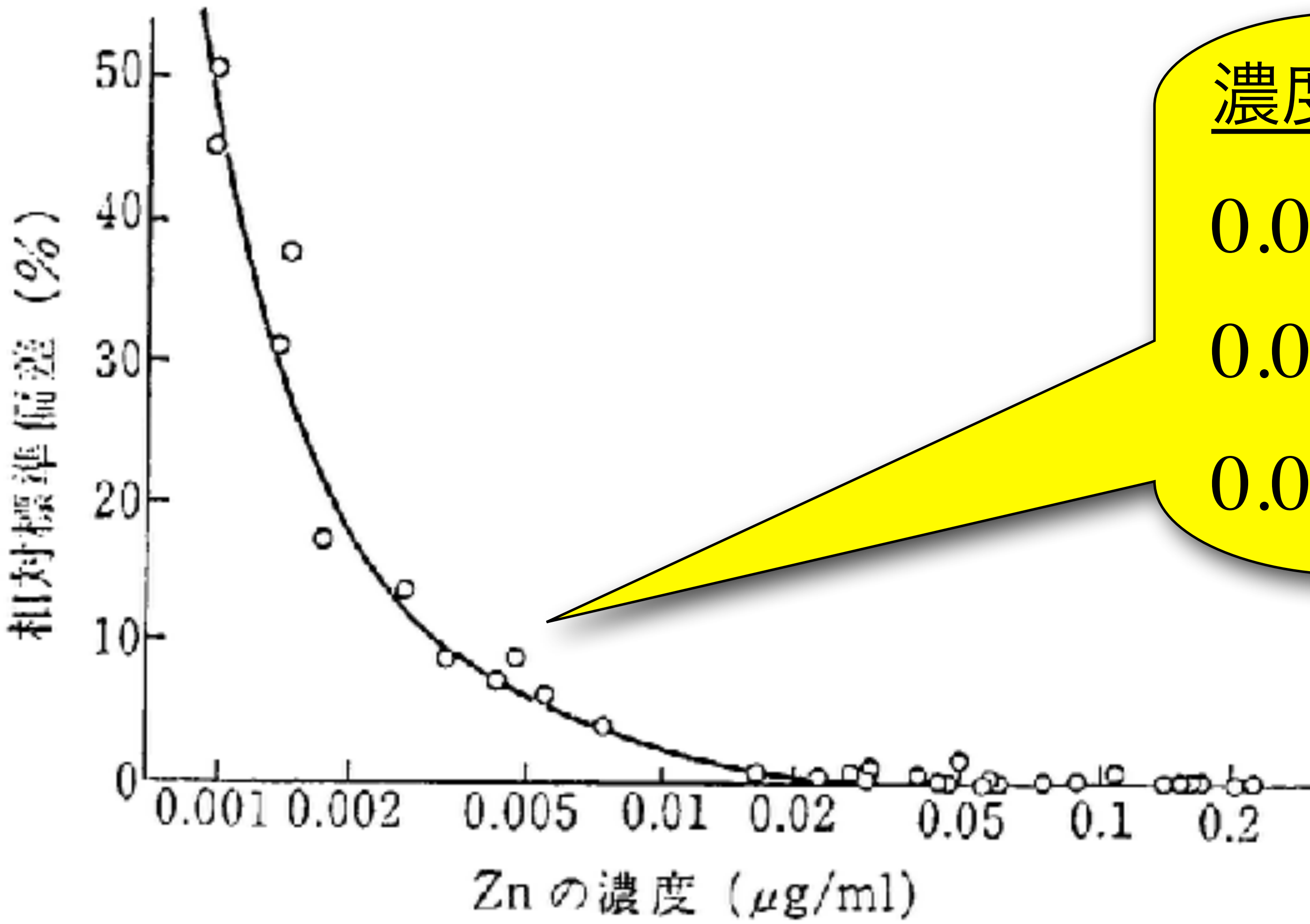


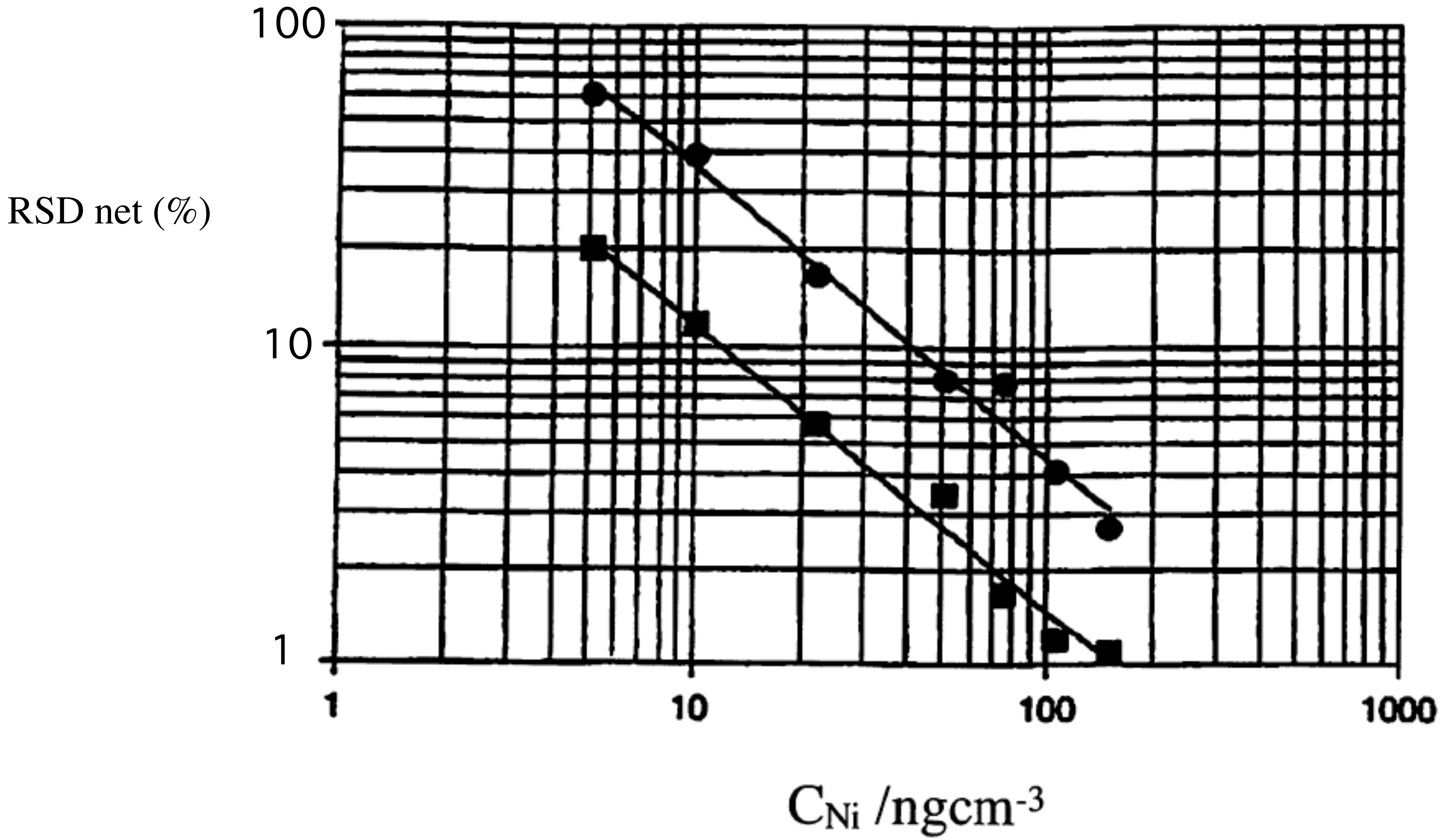
図８　相対標準偏差(RSD)と濃度の関係

濃度	相対標準偏差(%)	σ_Bのx倍
0.0009	～50	2
0.0045	～8	10
0.009	3～4	20

定量下限を一義的に10σ_Bの信号を与える濃度として定義することは正しくない

LOD（検出限界）= 0.0009 μgcm^{-3}

(Zn II 202.55 nm, 2σ_Bでの値)



ICP発光分析における実測値のRSDの濃度依存性

Ni(II 231 nm)の正味信号強度

積分時間;●1秒、■10秒、RSD5%の時に、1秒積分では90 ngcm^{-3}のLOQ値が、10秒積分では 25 ngcm^{-3}のLOQ値が得られる

検出限界や定量下限付近の分析値の考え方

事象	意味
検出限界(LOD)未満	検出せず
検出限界以上定量下限(LOQ)未満	検出された
定量下限以上	定量された

検出限界や定量下限付近の分析値の記載法

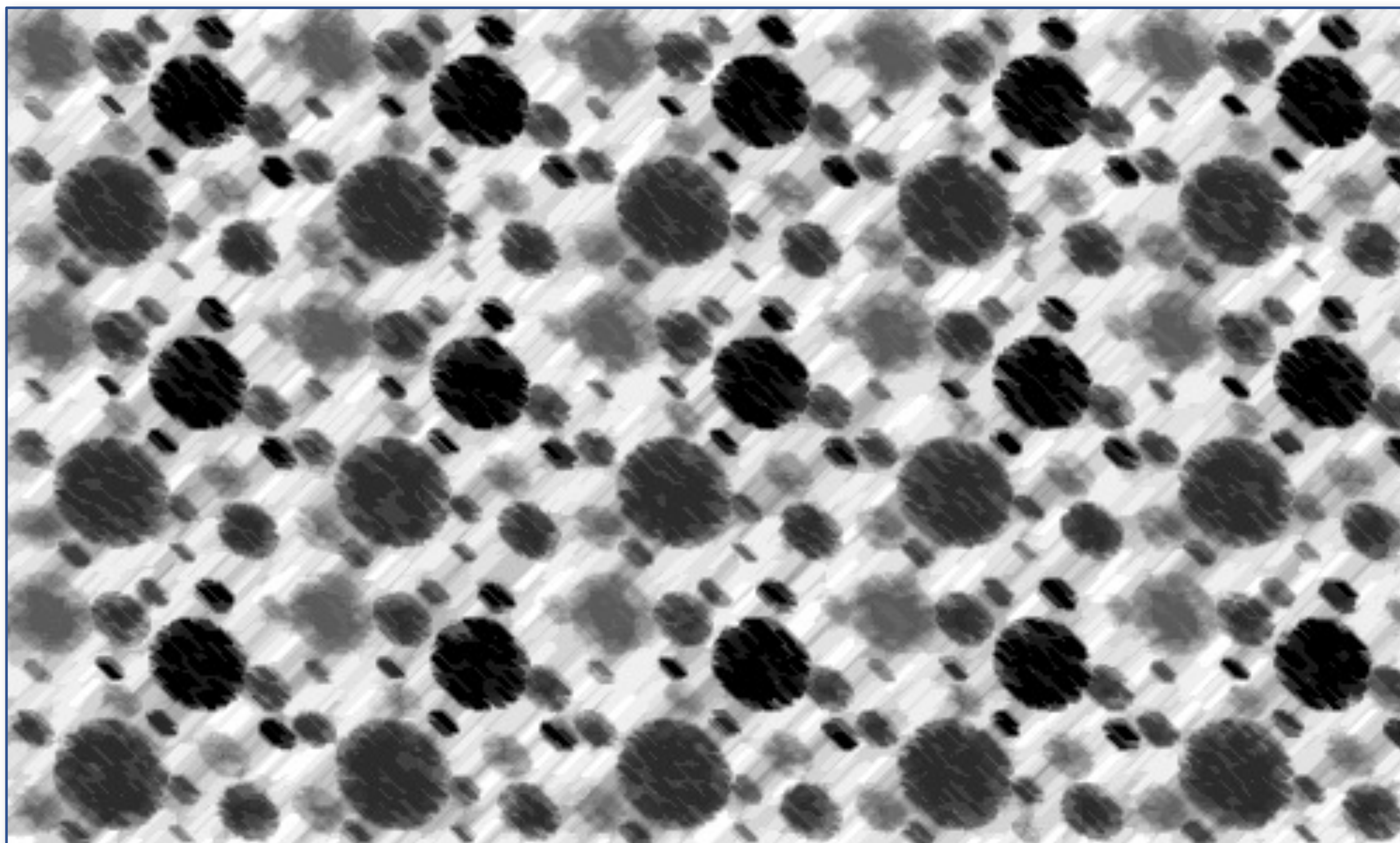
1）検出限界(LOD)をブランク値の3σ相当濃度として簡単のために3とする。それに対して測定結果が2であった場合、

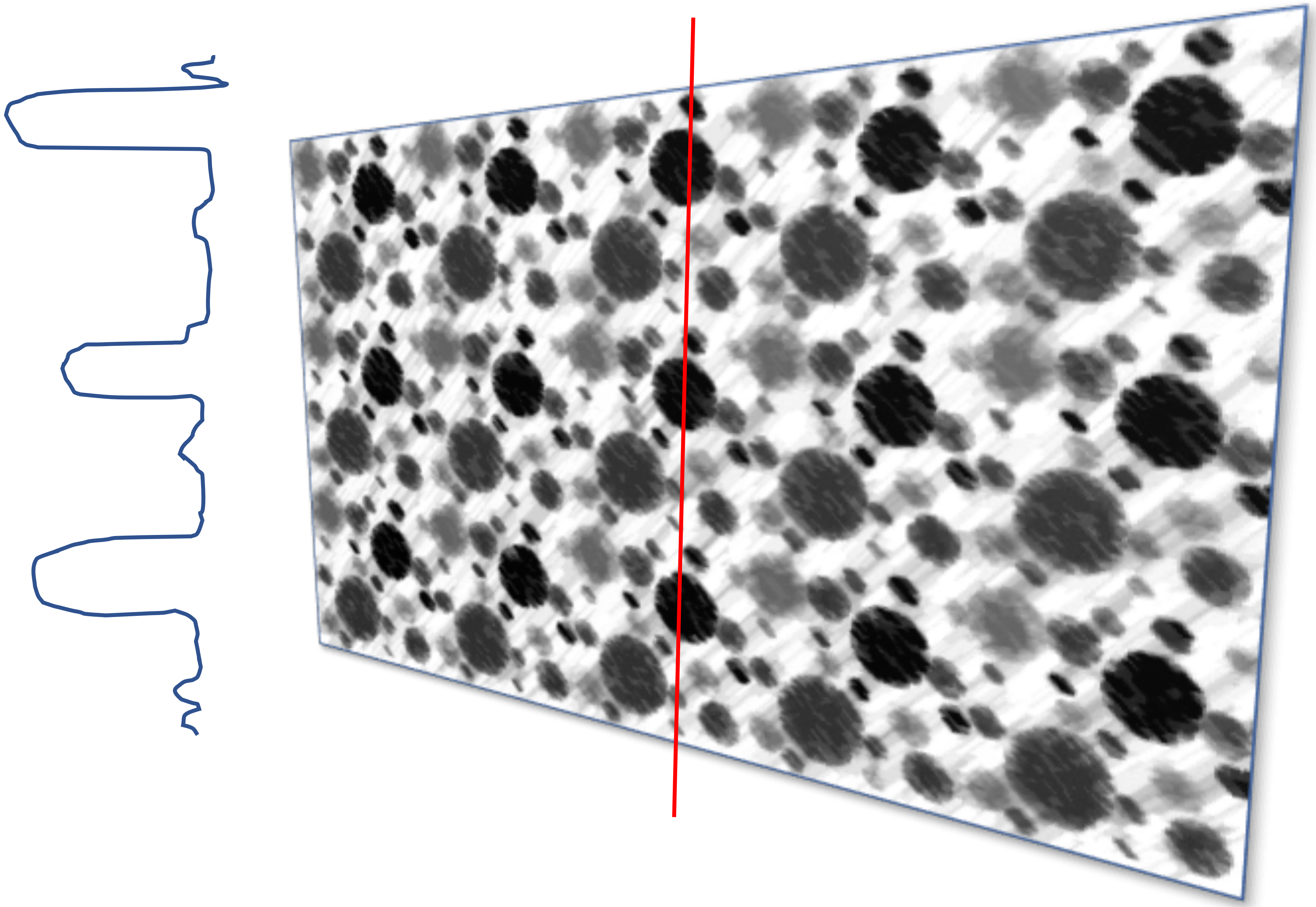
記載方法	記載例
検出限界(LOD)未満	<3
検出せず（LOD値を記載）	検出せず(LOD=3）

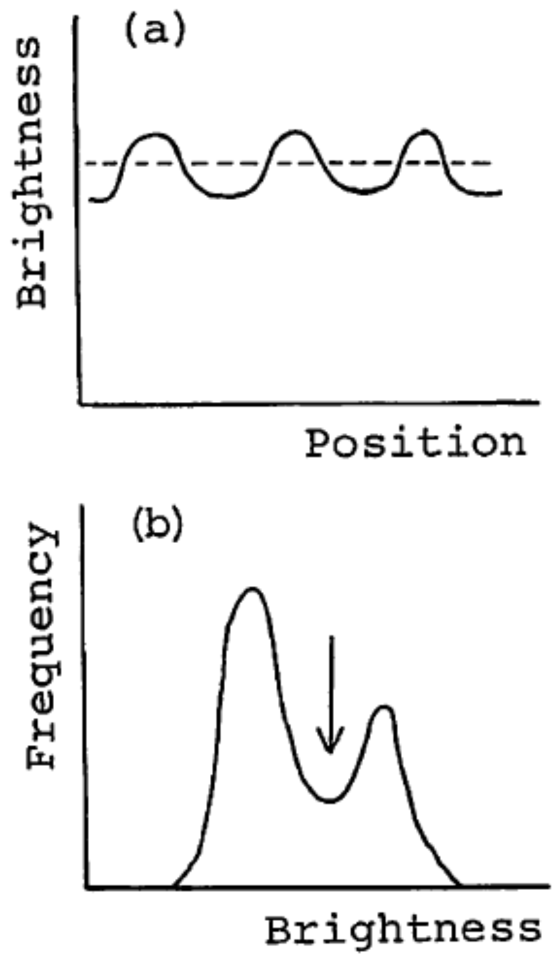
2）定量下限(LOQ)をブランク値の10σ相当濃度として簡単のために10とする。それに対して測定結果が6であった場合、

記載方法	記載例
検出限界以上で定量下限未満	>3であり<10
検出限界値と定量下限値の間	3から10
概ね6(定量下限値を記載)	概ね6（LOQ=10）
検出された(LODとLOQを記載）	検出された（LOD=3であり、LOQ=10）

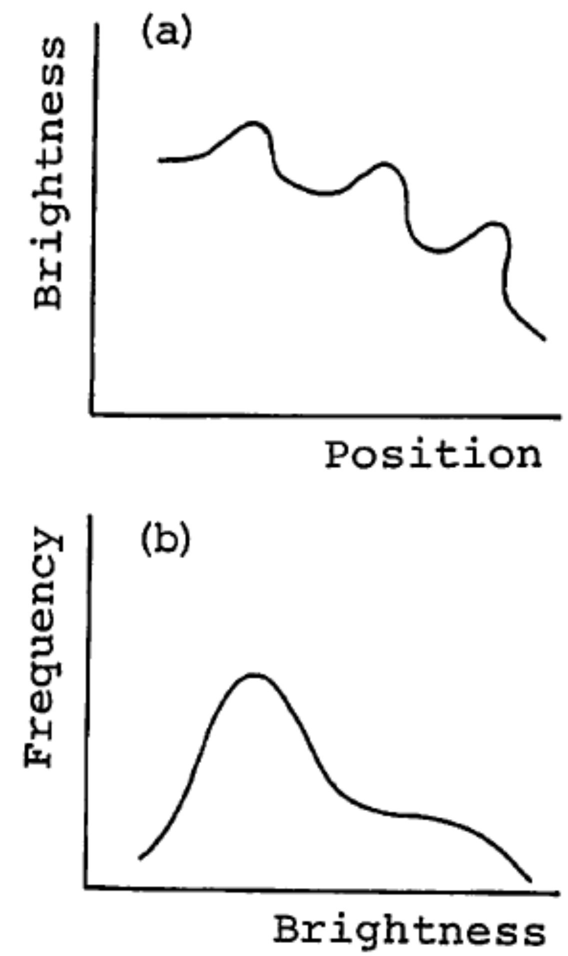
画像における検出の有効性をどう評価するか





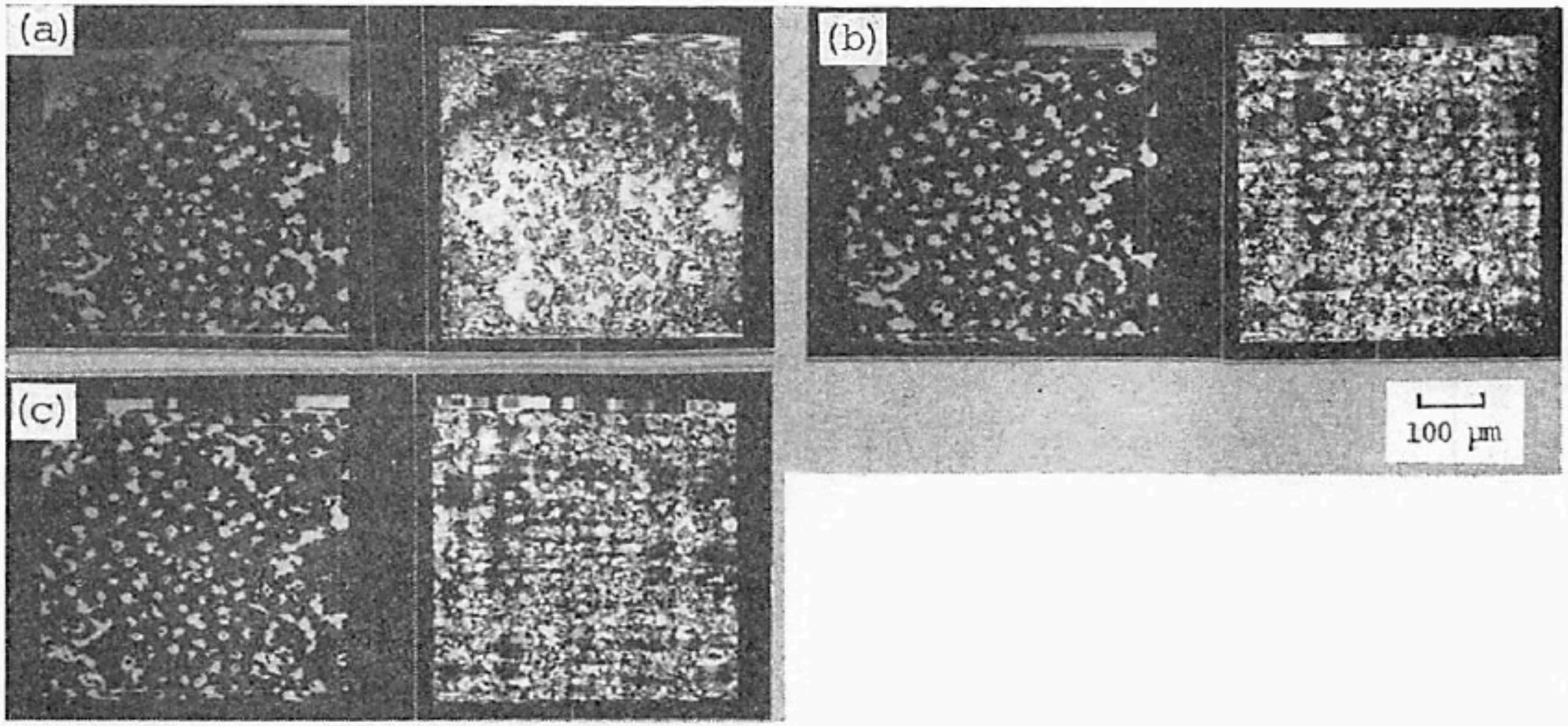


第２図　暗視野上の細胞の明るさの分布 (a) と明るさの頻度分布 (b)
(b) の矢印のスライスレベルが (a) の点線のレベルに相当する.



第３図　顕微鏡の視野に明るさむらがある場合の暗視野上の細胞の明るさの分布（ａ）と明るさの頻度分布（ｂ）

伊達、上本、貝原、田治：理研報告（1990）より



第４図　培養開始後２日めの BALB/c３T３細胞の顕微鏡写真を画像処理した結果
各写真の右はフーリエ変換によるフィルタリング処理を行った後のもの．左は二値化した結果の画像で，白い部分が細胞．(a) は高域（64次以上）カットのみ，(b) は高域および低域（１次）カット，(c) は高域および低域（１，２次）カットのフィルタリング処理をしたものである．

512 x 512 pixel, 8ビットカラーで画像取得

伊達、上本、貝原、田治：理研報告（1990）より

４．信頼性に関わる用語

不確かさ(uncertainty)；<u>測定結果に対する信頼度に関した、統一的な計測の信頼性表現</u>として、ISOを含む複数の国際機関が連携して提案した用語

不確かさの定義

- 「測定量の真の値が存在する範囲を示す推定値」

  JIS Z 8103「計測用語」(旧規格、1990)

VIM（International vocabulary of basic and general terms in metrology–国際計量基本用語集）第１版　に準拠

- 「測定の結果に付記される、合理的に測定量に結び付けられ得る値のばらつきを特徴づけるパラメーター」

  JIS Z 8103「計測用語」(2000改訂)

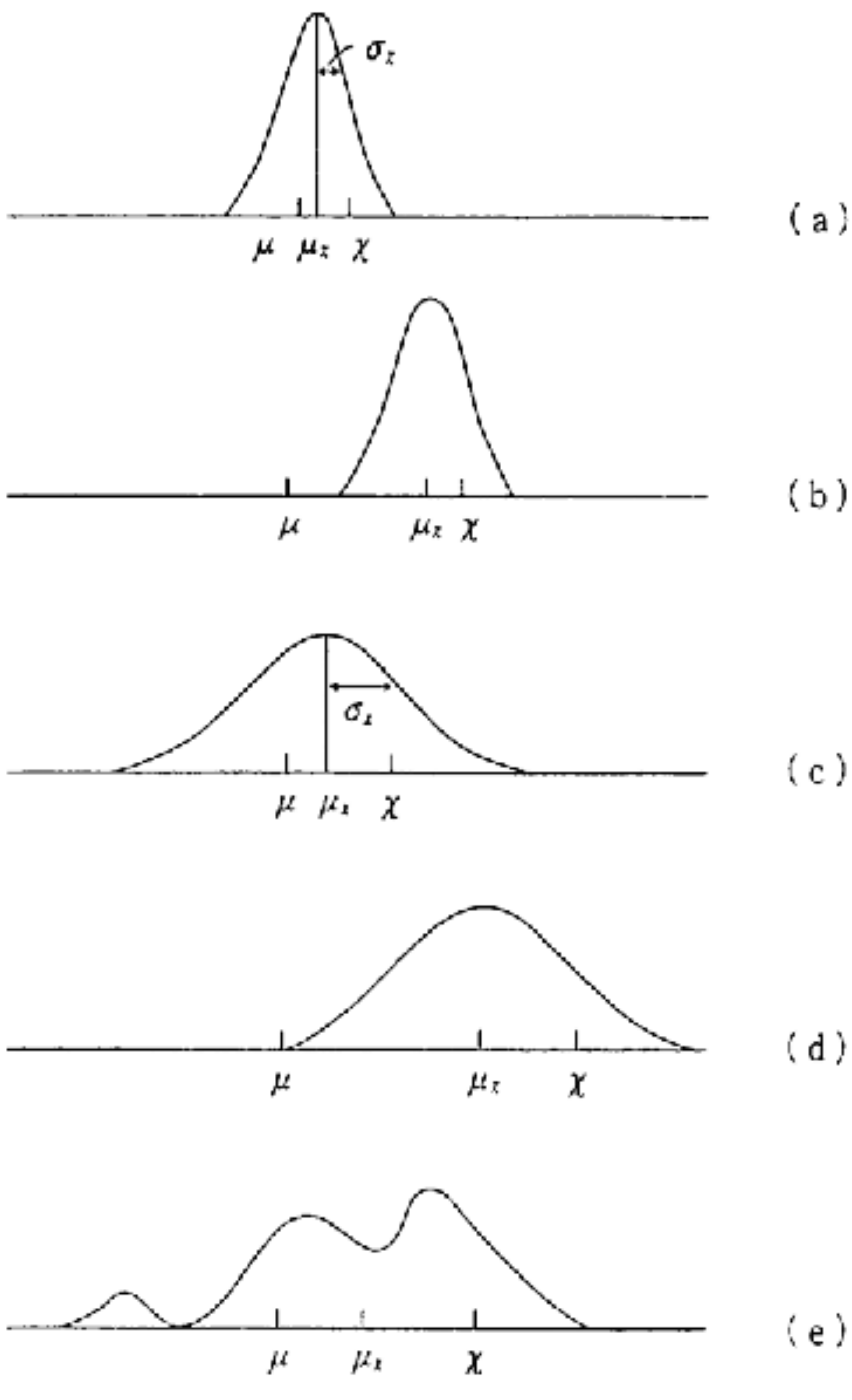
VIM第２版およびGUM（Guide to the expression of uncertainty in measurement–計測における不確かさの表現ガイド）第１版(1993)に準拠

信頼性の定義

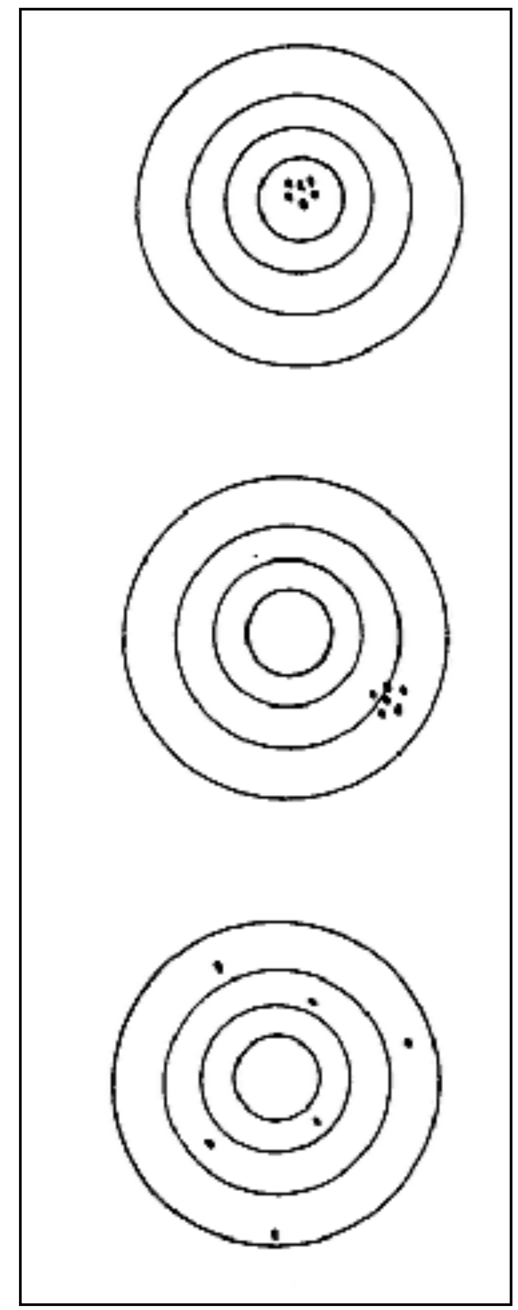
- 「精度又は正確さの期待できる程度」
  JIS K　0211「分析化学用語(基礎部門）」
  (旧規格、1987)

- 「機器、方法又はそれらの要素が、規定の条件の範囲内において規定の機能と性能を保持する性質又は度合い」
  JIS K　0211「分析化学用語(基礎部門）」
  (2013改訂)

データの分布状況と信頼性



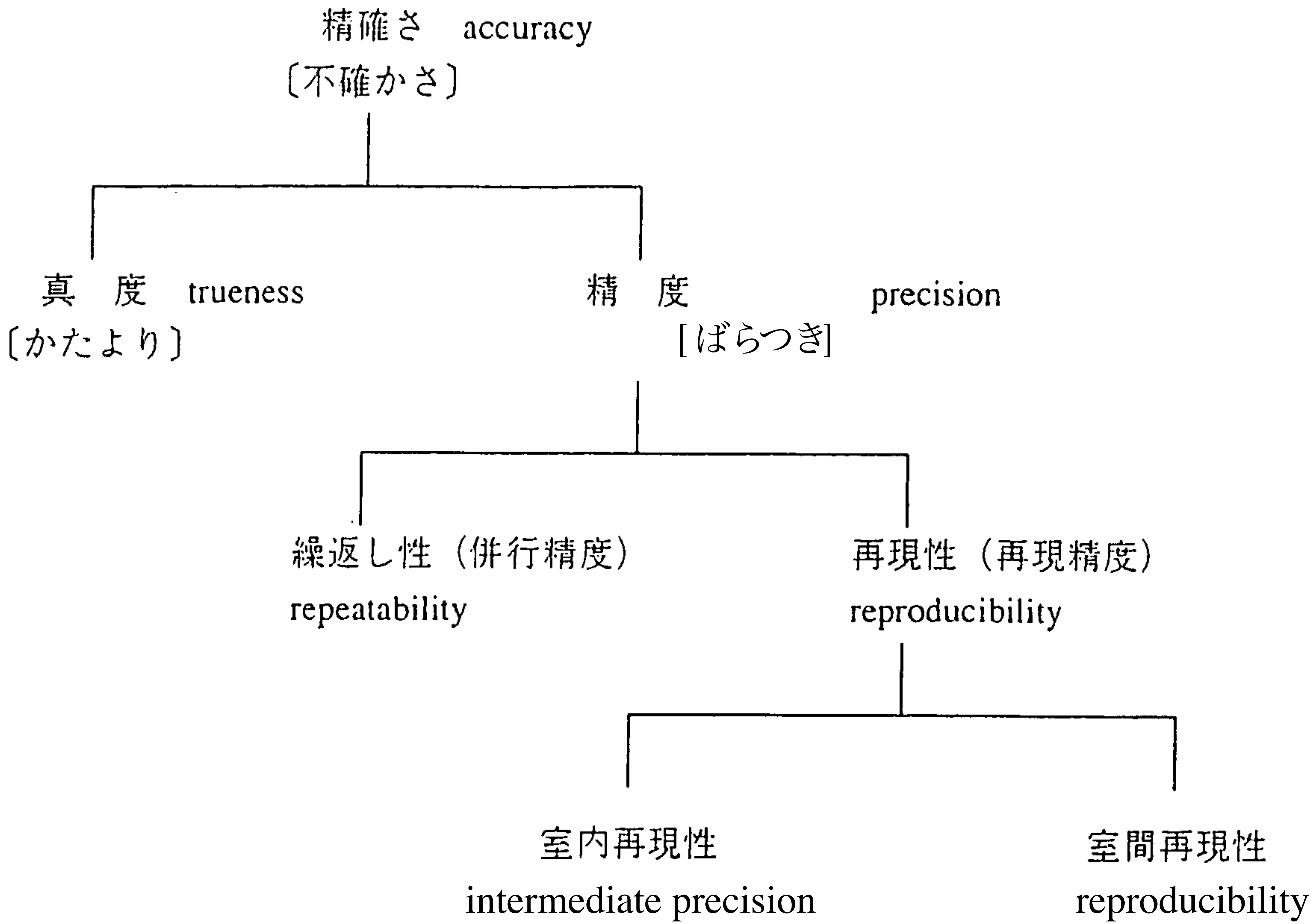
μ：真の値、μx：データの平均値、σx：標準偏差
○ 優れている、× 劣っている、− 求められない



True and precise

Precise, but not true

Neither precise nor true



化学計測領域における信頼性用語

基礎科学分野のJIS規格で規定される信頼性に関わる用語（1）

分野	JIS規格	対応ISO規格	accuracy	trueness	precision
化学計測	Z 8402-1: 1999	5725-1:1994	精確さ	真度	**精度**
				正確さ	
	K 0211: 2013	–	精確さ	真度	**精度**
物理計測	Z 8103: 2000	–	**精度**	正確さ	精密さ
					精密度
数理統計	Z 8101-2 :1999	3435-2:1993	精確さ	真度	精度
			総合精度	正確さ	精密度
					精確さ
適合性評価	Q 0033:2002	Guide 33: 2000	精確さ	真度	精度
			精度	正確さ	精密さ

基礎科学分野のJIS規格で規定される信頼性に関わる用語(2)

分野	JIS規格	対応ISO規格	repeatability	reproducibility
化学計測	Z 8402-1: 1999	5725-1:1994	併行精度	(室間）再現精度
			繰返し精度	
	K 0211: 2013	–	繰返し性	再現性
物理計測	Z 8103: 2000	–	繰返し性	再現性
数理統計	Z 8101-2 :1999	3435-2:1993	併行精度	(室間）再現精度
			繰返し精度	再現性
			繰返し性	
適合性評価	Q 0033:2002	Guide 33: 2000	併行精度	(室間）再現精度
			繰返し性	再現性

電子工業分野のJIS規格で規定される信頼性に関わる用語（1）

規格	用語	対応英語	定義（一部抜粋）
C 1002: 1975（電子計測器用語）	確度	accuracy limit of error	規定された状態において動作する機器の測定値又は供給値に対し、製造業者が明示した誤差の限界値。
	繰り返し性	repeatability	同一の方法で同一の測定対象を、同じ条件で比較的短い時間に繰り返し測定した場合の、個々の測定値が一致する度合。
	再設定性	resettability	供給量の設定において、同じ条件で比較的短い時間に、繰り返し同じ設定値を与えた場合の個々の供給値が一致する度合。
	安定性	stability	1　規定された時間、機器がその指示値、表示値、又は供給値を維持する能力。ただし、すべての状態を一定に保つ。 2　指定された一つの外部影響量のの変化に対して、機器がその指示値、表示値、又は供給値を維持する能力。ただし、他の状態を標準状態に保つ。
	測定値	measured value	測定によって求めた値
	指示値	indicated value	機器の指示値。アナログ機器の場合
	表示値	indicated value	機器の表示値。デジタル機器、オシロスコープ、記録計等の場合
	供給値	supplied value	機器が供給した量の値

電子工業分野のJIS規格で規定される信頼性に関わる用語(2)

規格	対応国際規格	用語	対応英語
B 0155: 1997 （工業プロセス計測制御用語及び定義）	IEC　902: 1987	正確さ	accuracy
		組合せ精度 総合精度	system accuracy
		精度定格	accuracy rating
		一致性	conformity
		独立一致性	independent conformity
		ヒステリシス	hysteresis
		不感帯	dead band
		ドリフト	drift
		繰返し性誤差	repeatability error
		再現性誤差	reproducibility error

JIS C1002における、規定された時間中の機器の指示値、表示値若しくは供給値の変化度合

ドリフト	drift	穏やかで継続的な、好ましくない変化
PARD	periodic and/or random deviation	平均値付近における、周期的若しくはランダムな、又はその双方を含む好ましくない変化
ハム	hum	平均値付近における、電源周波数に関連したほぼ正弦波状の低周波の好ましくない変化
リプル	ripple	平均値付近における、周期的であるが非正弦波状の好ましくない変化
雑音	noise	1　広い周波数範囲にわたり、ランダムに生じる好ましくない変化 2　信号に重畳し（かさなり）、あいまいにする妨害。
揺らぎ	fluctuation	平均値付近における、ランダムで比較的ゆっくりした好ましくない非周期的変化

精度(*precision*)と対応をつけるべき語である！

信頼性に関わる用語が規定されているJIS

＜基礎科学分野＞

- 化学計測；K 0211:2013　分析化学用語（基礎部門），
　Z 8402-1 :1999　測定方法及び測定結果の精確さ（真度及び精度）–第１部：一般的な原理及び定義
- 物理計測；Z 8103 : 2000　計測用語
- 数理統計； Z 8101-2 :1999　統計–用語と記号–第２部：統計的品質管理用語
- 適合性評価；Q0033:2002　認証標準物質の使い方

＜応用分野＞

- 電子工業；C1002: 1975　電子計測器用語
- 工業プロセス；B0155:1997　工業プロセス計測制御用語及び定義

おわりに

- 分析をはじめ、あらゆる計測の際は、測定値の信頼性をいつも意識する。
- 検出限界と定量下限は、測定者が分析の目的を考えて実験的かつ合理的に決める。
- 信頼性用語を使う際は、分野と根拠規格を明示した方が賢明である。
- 測定値を元に分析値を提示するのは分析者の責任である。

michihisa.uemoto@meisei-u.ac.jp

検定の考え方と実際

京都大学　大学院工学研究科
西　直哉

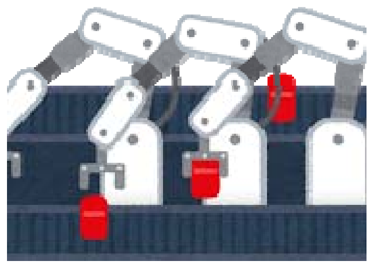
検定

偶然なのか、そうじゃないのか？

データは、ばらつく。
ばらつきを超えて差があること(＝有意差)を主張できるか？

検定の例

- 新しく開発した薬に、従来品を上回る効果があるか？

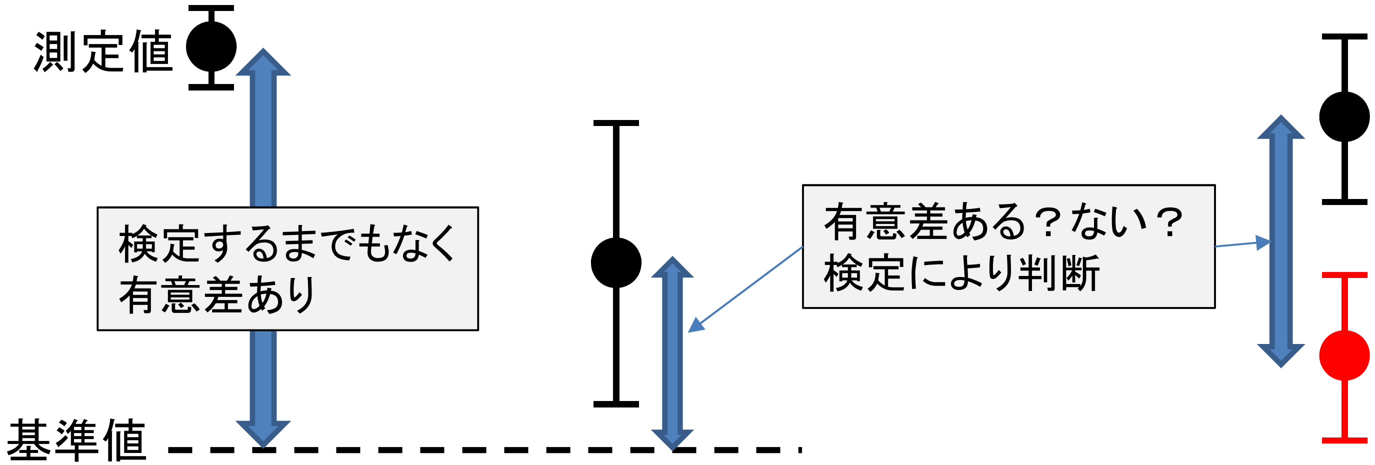


- 製品Aと製品Bの品質に違いはあるか？
- AさんとBさんの測定値および測定精度に違いはあるか？
- このサイコロはいかさまではないのか？

統計的推測ー検定ー

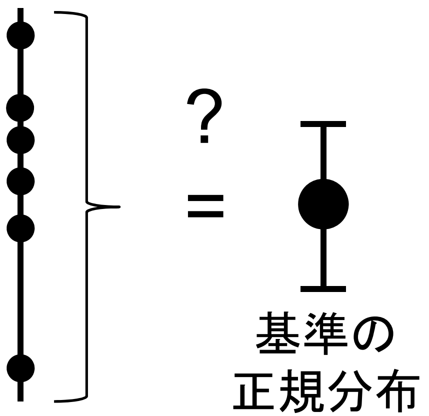
標本平均の検定

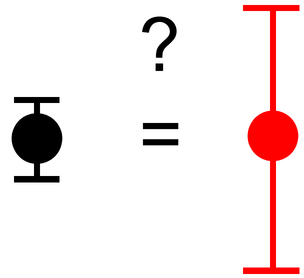
- 測定値と基準値に有意差があるか？
- 二つの測定値に有意差があるか？



分布・分散の検定

- データは基準の正規分布に従うか？
- 二つの精度に違いはあるか？





各種検定法

- 標本平均の検定
  - 測定値と基準値（あるいは二つの測定値）に有意差があるか？

    *z*検定（母分散が既知）

    *t*検定（母分散が未知）

- 分布・分散の検定
  - 測定値のばらつき方は基準の正規分布に等しいか？

    χ^2検定　（χ^2分布は、正規分布に従う変数の分散の分布）
  - 二つの測定値の分散に差があるか？

    *F*検定　（*F*分布は、χ^2分布に従う二つの変数比が示す分布）

- 3つ以上のデータ群の平均値の差、2つ以上の条件が掛け合わさったデータ間の相関

  分散分析 (ANalysis Of VAriance、ANOVA)

- 外れ値の棄却

  Dixon法（Qテスト）、Grubbs法、Hampel法など

検定の流れ

1. その問題に適切な検定法（および統計量 X ）を決める

   z検定、t検定、χ^2検定、F検定・・・

2. 対立仮説H_1／帰無仮説H_0を決める

   対立仮説により両側検定か片側検定かが決まる

3. 有意水準 α を決め、臨界値 X_α を計算する

   α = 5%とすることが多い　　信頼度（1-α）は95%

4. 帰無仮説H_0を仮定して、統計量 X を計算する

   偶然と仮定すると、どれくらいあり得ない事象か？

5. 判定　　XはX_αよりあり得ない値か？

   yes　→　対立仮説H_1を採用する　　偶然ではない

   no　→　帰無仮説H_0は棄却できない　偶然ではないとは言えない

本講習の内容

本講習のゴール：二つの測定値の有意差検定

【例題】下表は、同じ溶液中のある化学種の濃度を、A研究室とB研究室がそれぞれ独自に測定した結果である。

	平均 (ppm)	標準偏差 (ppm)	試行回数
A研究室	$\bar{x}_1 = 7.85$	$u_1 = 0.61$	5
B研究室	$\bar{x}_2 = 6.34$	$u_2 = 0.83$	7

測定値に有意差があるかどうかを有意水準5%で検定せよ。

本講習のマインドマップ

	分散の検定		標本平均の検定
例題でやること	F検定⑥	⇒	t検定（二つの測定値）④
理解を深めるためにやること	χ^2検定⑤		・z検定（両側検定）①講習で出てくる順番 ・z検定（片側検定）② ・t検定（基準値に対する）③

z検定：測定値と基準値の比較

測定値は基準値に対して有意差があるか？
（母分散が既知）

【問題0-1】

工場のラインにおいて、製品の品質を反映する物理量 x を自動測定して、品質管理をしている。x の自動測定結果は、母分散 $\sigma_0^2 = 10^2$の正規分布に従うことがわかっている。正常な製品の x は母平均 $\mu_0 = 50$を示す。

あるロットの製品4個が、標本平均 $\bar{x} = 58$を示した。このロットの製品は異常だろうか？　有意水準5%で判断せよ。

流れ1．適切な検定を決める

統計量は正規分布に従う　→　標準正規分布に従うzを用いる　→　z検定

帰無仮説と対立仮説

偶然なのか、そうじゃないのか？

偶然ではない → 対立仮説H_1
偶然だ → 帰無仮説H_0 → 棄却できるか検定する

流れ2．対立仮説／帰無仮説を決める

問題0-1においては
対立仮説H_1：測定値は基準値に対して有意差がある $(\mathrm{E}[\bar{x}] = \mu_1 \neq \mu_0)$
帰無仮説H_0： 〃 ない $(\mathrm{E}[\bar{x}] = \mu_1 = \mu_0)$

有意差のあるなしなので両側検定 （両側検定・片側検定については後述）

H_0が成り立つなら、測定値xの母平均μ_1は基準値μ_0と等しいはず

$$\boxed{z = \frac{\bar{x} - \mu_0}{\sigma_0 / \sqrt{N}}} \Rightarrow$$ 標準正規分布に従うはず

z検定：有意水準と臨界値

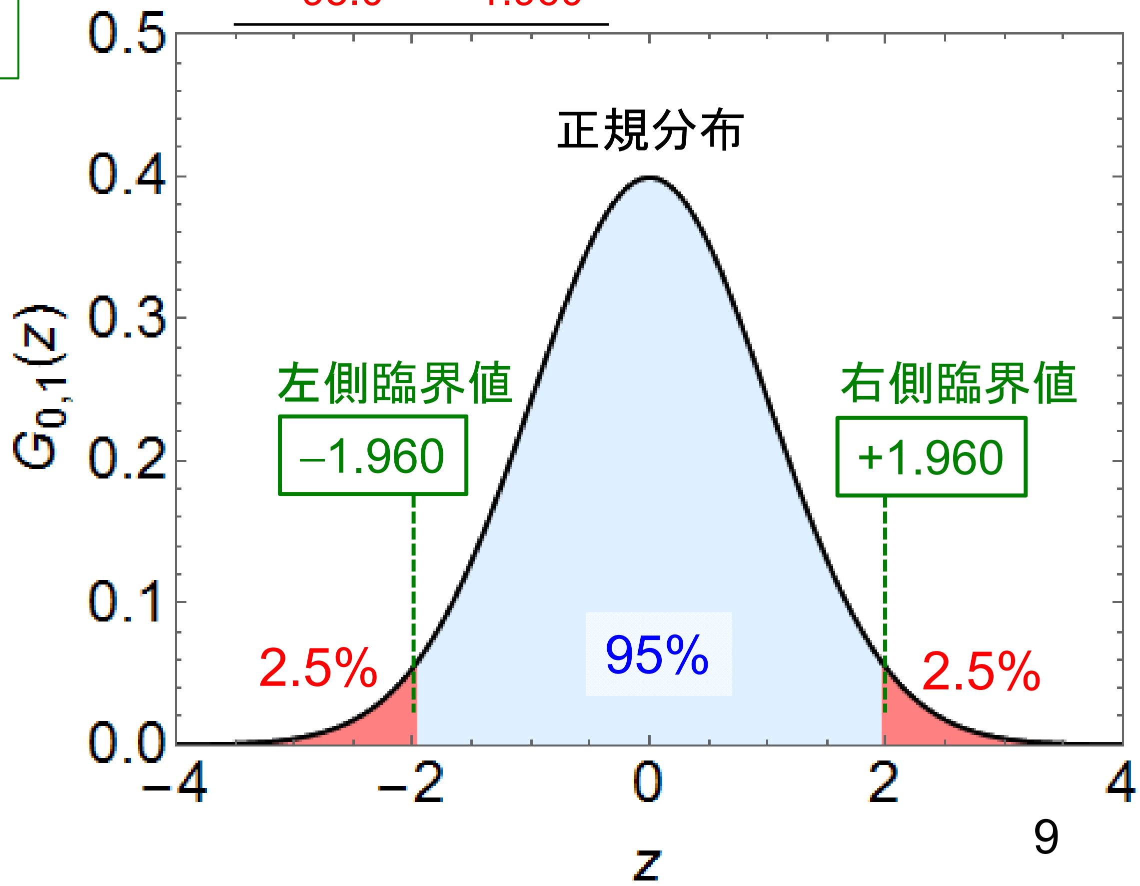
流れ3．有意水準 α を決め、臨界値 z_α を計算する

$$z_{5\%} = \pm 1.960$$

$(1-\alpha)$ / %	z_α
68.3	1.000
90.0	1.645
95.0	1.960

測定回数 N=4
有意水準　α=5%
信頼度　1-α=95%

ExcelまたはGoogle Sheetsで
“NORM.S.INV(0.025)”
で左側臨界値が求められます



右図の
青：「偶然であり得るゾーン」
赤：「偶然であり得ないゾーン」

値の小さいほうを左側
　　大きいほうを右側
と呼びます
（臨界値や片側検定）

z検定：判定

流れ4．帰無仮説H_0を仮定して統計量を計算する

測定回数 N=4
有意水準　α=5%
信頼度　1-α=95%

$$z = \frac{\bar{x} - \mu_0}{\sigma_0 / \sqrt{4}} = \frac{58 - 50}{5} \quad = 1.6$$

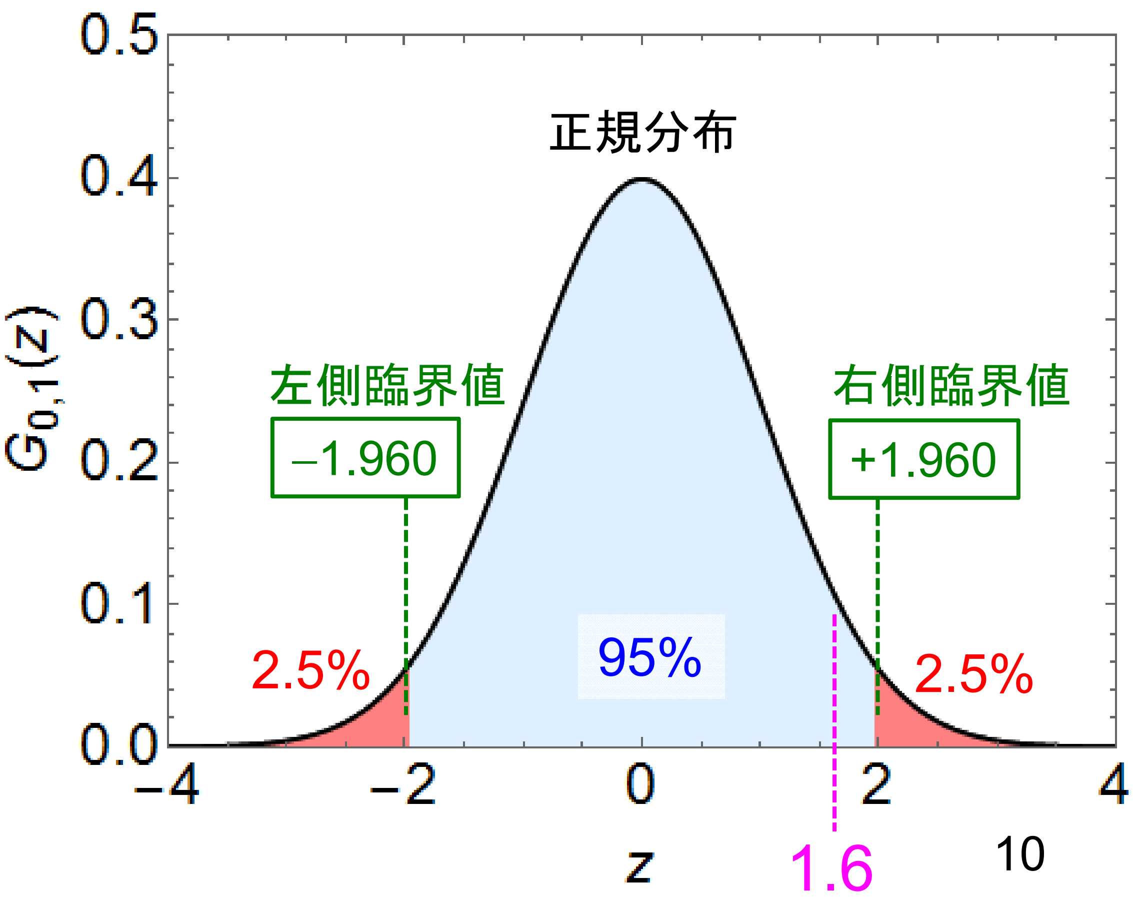
流れ5．判定

$-1.960 < 1.6 < 1.960$

この測定値を測定する確率は95%の範囲内に入っており、その測定結果は偶然に起こりえたと言える

H_0は棄却できない

有意水準5%でこのロットの製品が異常とは言えない



z検定II：片側検定の例

【問題0-2】

工場のラインにおいて、製品の品質を反映する物理量 x を自動測定して、品質管理をしている。x の自動測定結果は、母分散 $\sigma_0^2 = 10^2$の正規分布に従うことがわかっている。正常な製品の x は母平均 $\mu_0 = 50$を示す。

この x は大きい場合には問題ないが、小さい場合には異常とみなし、その製品の出荷を見合わせないといけない。あるロットの製品4個が、標本平均 $\bar{x} = 41$を示した。このロットの製品は異常だろうか？　有意水準5%で判断せよ。

1．適切な検定を決める
統計量は正規分布に従う　→　標準正規分布に従うzを用いる　→　z検定

両側検定と片側検定

流れ2．対立仮説／帰無仮説を決める

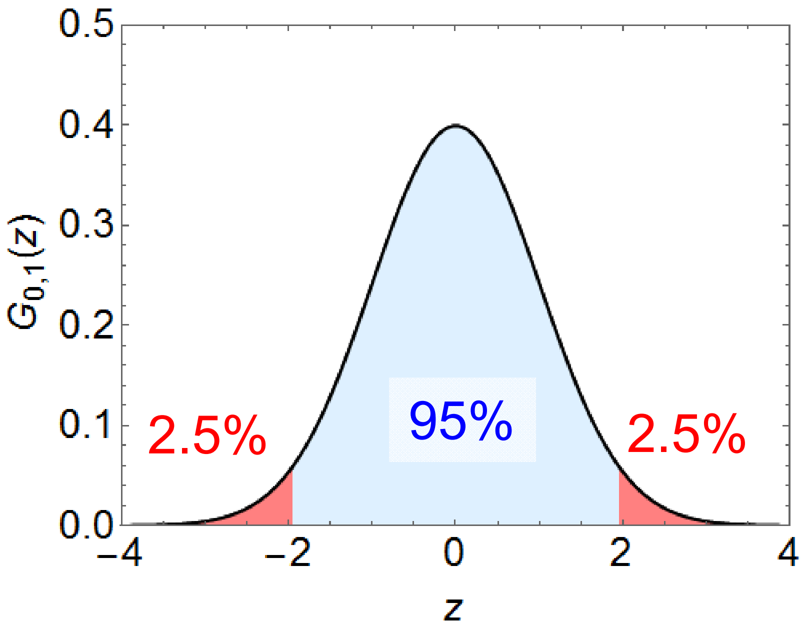
【問題0-1】

測定値は基準値と異なるか？

両側検定

対立仮説H_1　$\mu_1 \neq \mu_0$

帰無仮説H_0　$\mu_1 = \mu_0$



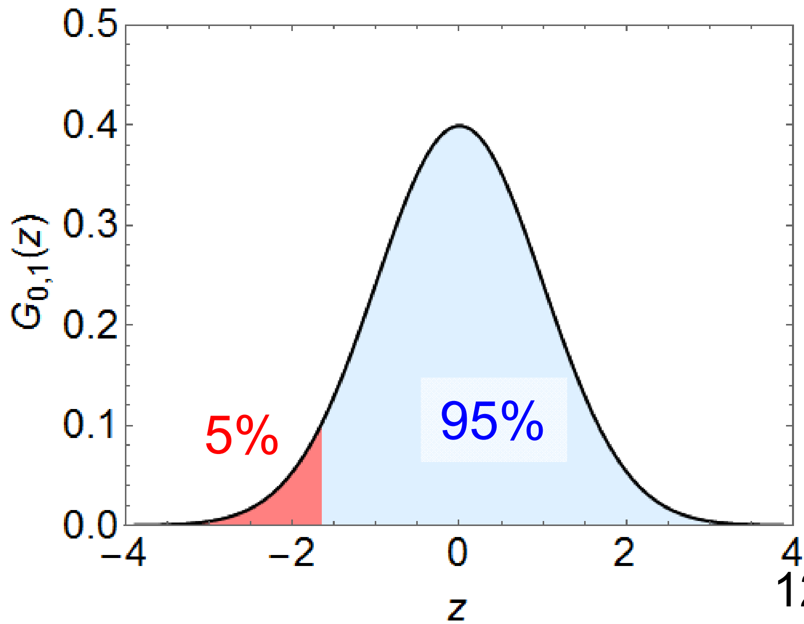
【問題0-2】

測定値は基準値より小さいか？

片側検定（左側）

対立仮説H_1　$\mu_1 < \mu_0$

帰無仮説H_0　$\mu_1 = \mu_0$



z検定II：有意水準と臨界値

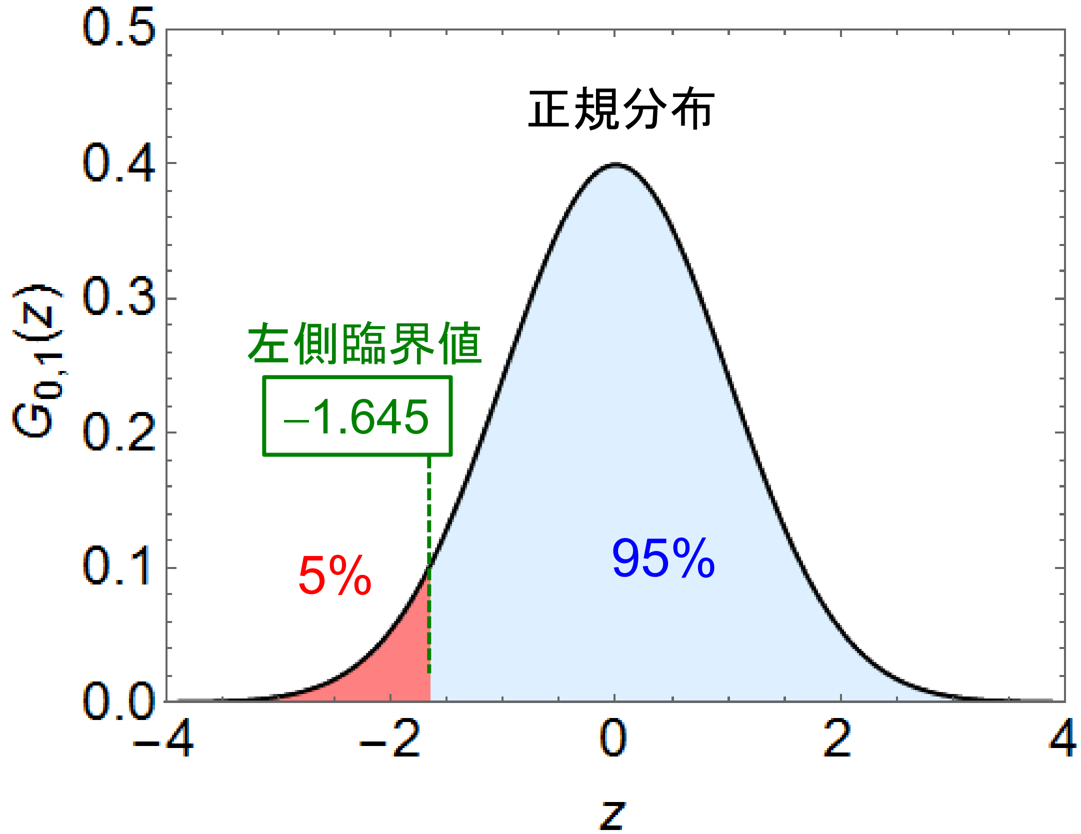
流れ3．有意水準 α を決め、臨界値 z_α を計算する

測定回数 N=4
有意水準　α=5%
信頼度　1-α=95%

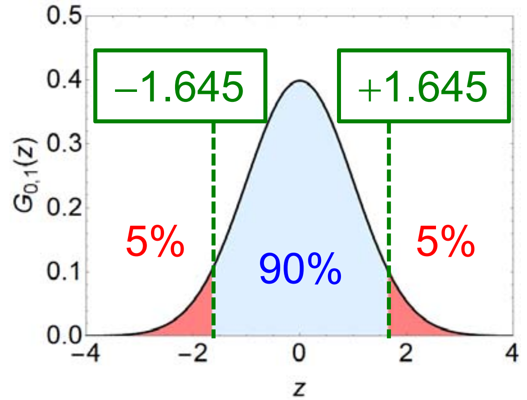
$$z_{5\%,\mathrm{left}} = -1.645$$

“NORM. INV(0.05,0,1)”

$(1-\alpha)$ / %	z_α
68.3	1.000
90.0	1.645
95.0	1.960



$$z_{10\%} = \pm 1.645$$



z検定II：判定

流れ4．帰無仮説を仮定して統計量を計算する

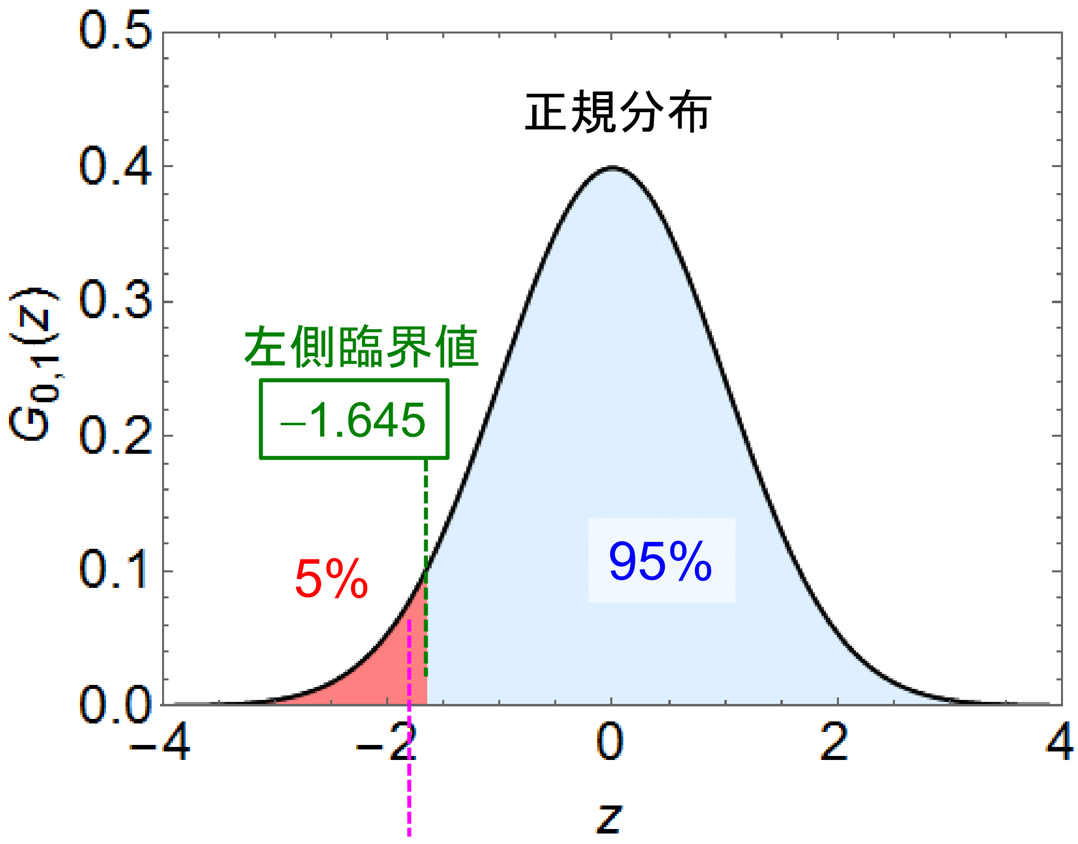
測定回数 N=4
有意水準 α=5%
信頼度 1-α=95%

$$z = \frac{\bar{x} - \mu_0}{\sigma_0 / \sqrt{4}} = \frac{41 - 50}{5} = -1.8$$

流れ5．判定

$-1.8 < -1.645$
95%の範囲外にあり、偶然とはいえないので、H_0を棄却し、H_1を採用する。

有意水準5%で、このロットの製品は異常といえる。



$-1.8 < -1.645$

t検定

測定値は基準値に対して有意差があるか？
（母分散は未知）

【問題1】

Mgの当量を4回測定し、以下の測定値を得た。

標本平均 $\bar{x}$ = 12.22、不偏分散 $u^2 = 0.05^2$

原子量からMgの当量を計算すると、基準値μ_0 = 12.15 である。
測定値は基準値に対して有意差があるか？（測定に系統誤差があるか？）有意水準5%で判断せよ。

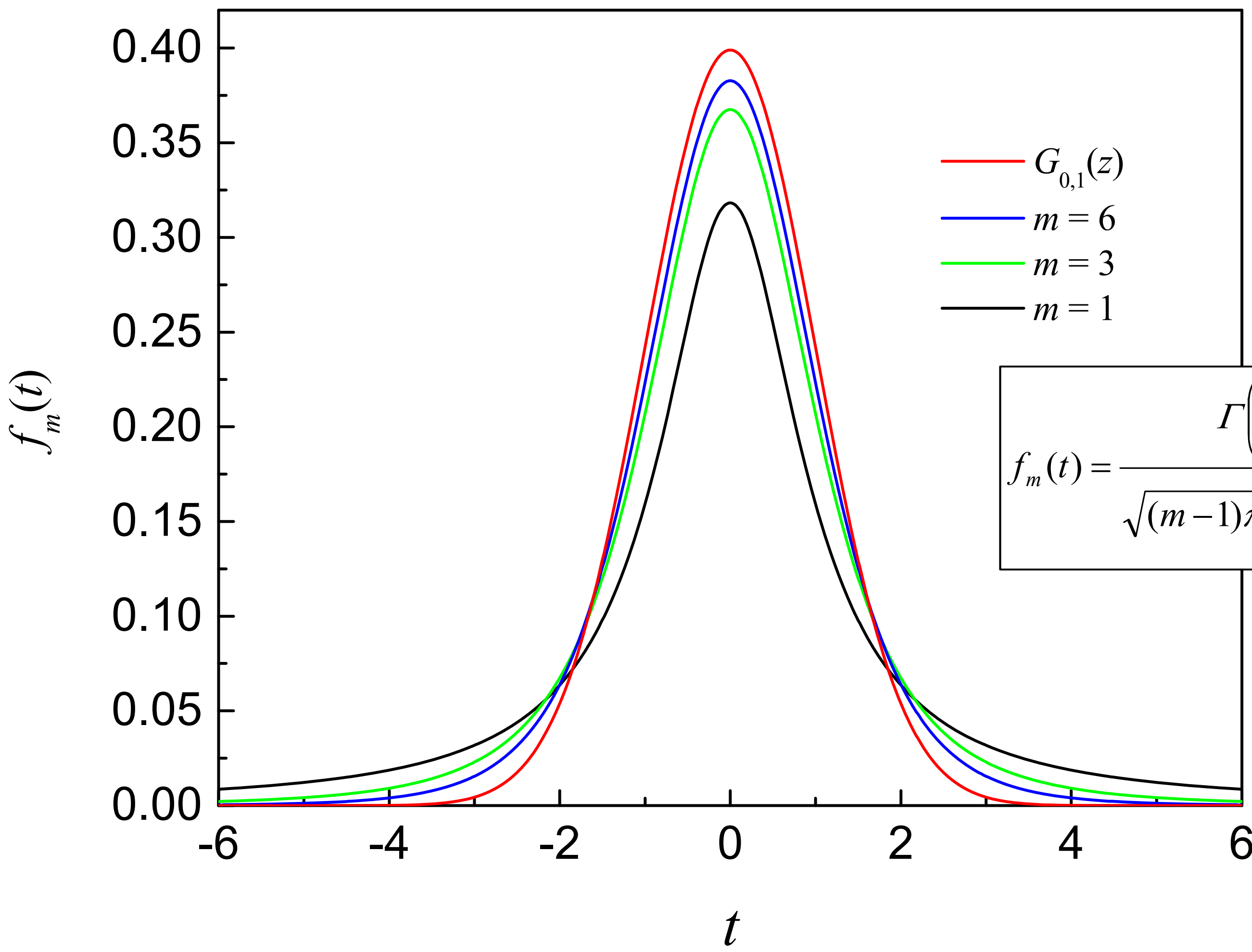
1．適切な検定を決める

測定精度は未知 → 不偏分散から母分散を推定するt分布 → t検定

（もし測定精度が既知なら → 正規分布 → z検定）

Studentのt-分布

測定値xの母平均 = μ_1　　$\boxed{t = \dfrac{\bar{x} - \mu_1}{u/\sqrt{n}}}$ ← t分布に従う



$$f_m(t) = \frac{\Gamma\left(\frac{m}{2}\right)}{\sqrt{(m-1)\pi}\,\Gamma\left(\frac{m-1}{2}\right)}\left(1 + \frac{t^2}{m-1}\right)^{-m/2}$$

m:自由度

$m = n-1$

t検定：帰無仮説と対立仮説

流れ2．対立仮説／帰無仮説を決める

対立仮説 H_1：測定値は基準値に対して有意差がある　$(\mathrm{E}[\bar{x}] = \mu_1 \neq \mu_0)$
帰無仮説 H_0：　　　　〃　　　　　　　ない　$(\mathrm{E}[\bar{x}] = \mu_1 = \mu_0)$

有意差のあるなしが判定基準なので、両側検定

H_0が成り立つなら、測定値xの母平均μ_1は基準値μ_0と等しいはず

$$\boxed{t = \frac{\bar{x} - \mu_0}{u / \sqrt{n}}} \Rightarrow \text{ t分布に従うはず}$$

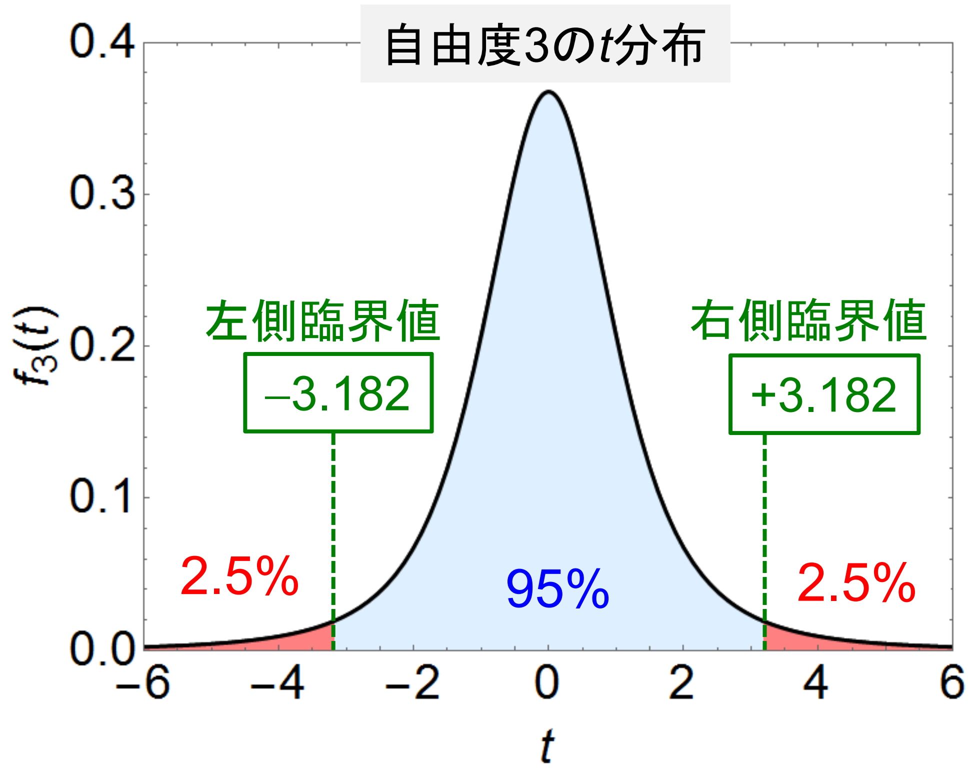
t検定：有意水準と臨界値

流れ3．有意水準 α を決め、臨界値 $t_{N\text{-}1,\alpha}$ を計算する

測定回数 N=4
有意水準　α=5%
信頼度　1-α=95%

$$t_{4-1,5\%} = \pm 3.182$$

"T.INV.2T(0.05,3)"



$z_{5\%} = \pm 1.960$ よりも絶対値が大きい。分布が広がっているため。

t検定：判定

流れ4．帰無仮説を仮定して統計量Xを計算する

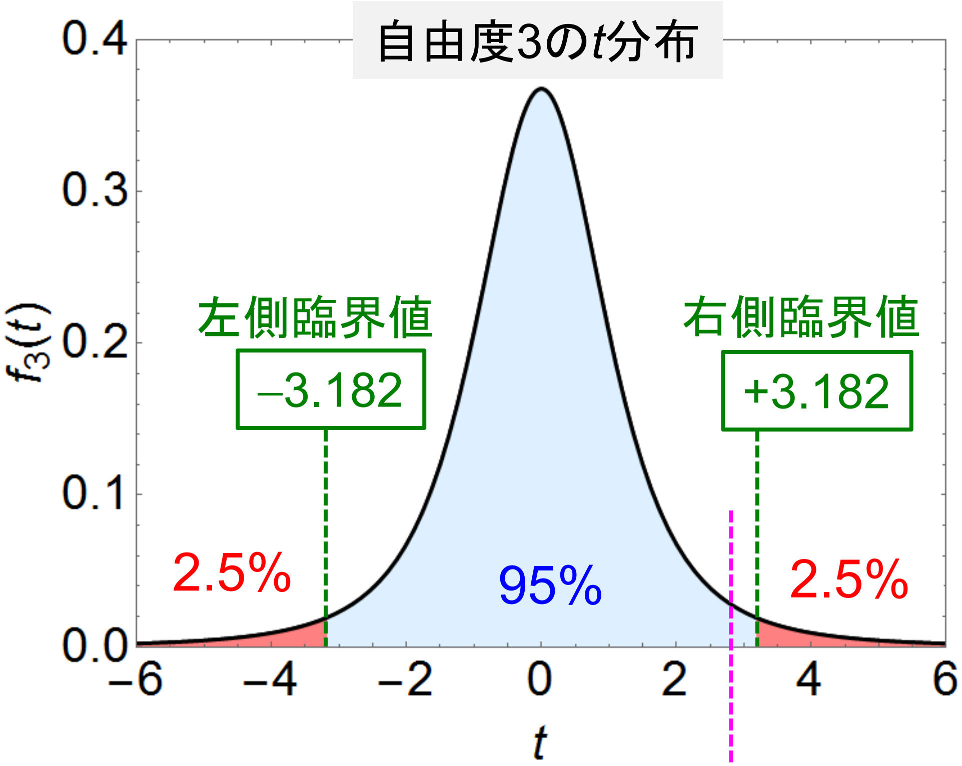
測定回数 N=4
有意水準 α=5%
信頼度 1-α=95%

$$t = \frac{\bar{x}-\mu_0}{u/\sqrt{n}} = \frac{12.22-12.15}{0.05/\sqrt{4}} = 2.8$$

流れ5．判定

$-3.182 < 2.8 < 3.182$
よりH_0は棄却できない

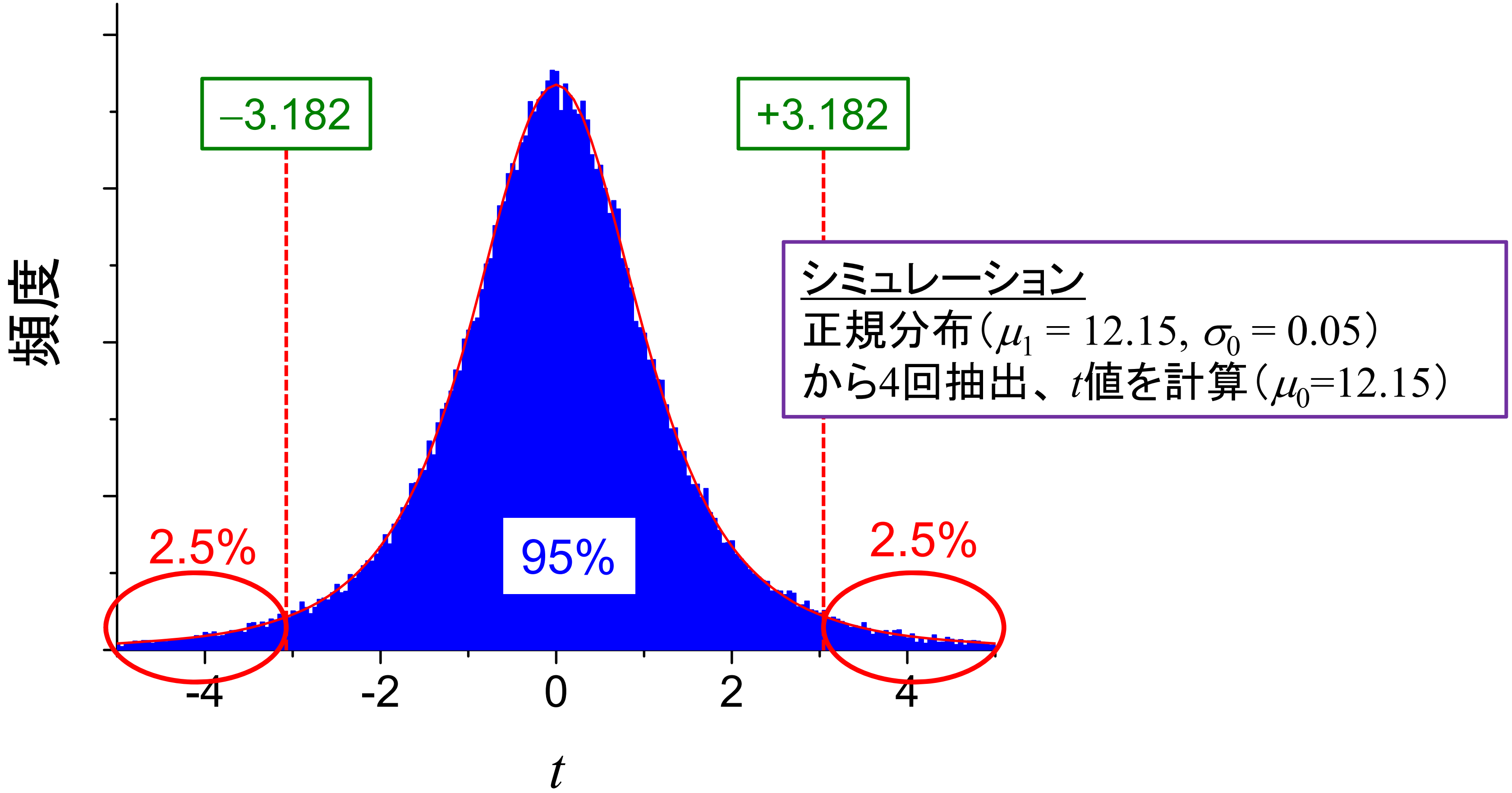
Mg当量の測定値は基準値に対して有意差があるとはいえない



$-3.182 < 2.8 < 3.182$

帰無仮説H_0が真　$\mu_1 = \mu_0$

✓ 帰無仮説H_0が正しい場合　⇒　t分布に従う



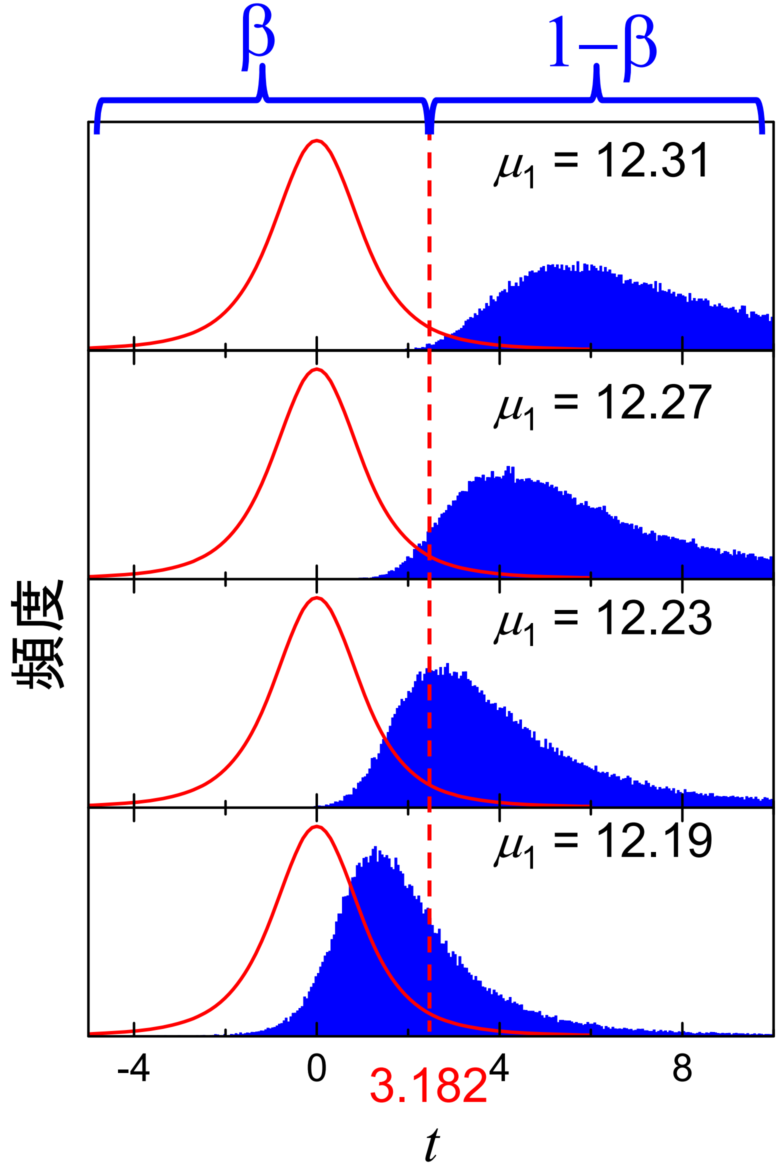
H_0が正しくても棄却してしまう誤り…第一種の誤り

有意水準（危険率）α = 第一種の誤りが起こる確率

「あわてもの」の誤り

帰無仮説H_0が偽 $\mu_1 \neq \mu_0$

✓ 帰無仮説H_0が誤っている場合 ⇒ t分布からずれる



シミュレーション
正規分布(μ_1, σ_0 = 0.05)
から4回抽出、t値を計算(μ_0=12.15)

μ_1とμ_0の差：大
⇒ t値の平均値が0からずれる
分布の形も変わる

H_0が偽でも棄却しない誤り
…第二種の誤り

第二種の誤りが起こる確率=β

「ぼんやりもの」の誤り

「あわてもの」と「ぼんやりもの」

		検定結果	
		H_0を棄却できない	H_0を棄却し H_1を採用
事実	H_0が真	正しい $1-\alpha$	第一種の誤り α（あわて）
	H_1が真	第二種の誤り β（ぼんやり）	正しい $1-\beta$

- H_1を誤って採用する（あわてる）確率αは設定により制御可能
  α=5%に設定することが多い

- H_1を見逃す（ぼんやりする）確率βは直接は制御できない
  βは
  - n　測定回数
  - $\mu_1 - \mu_0$　真値と基準値のずれ
  - α　有意差

  により決まる

αはなぜ5%？

Q.有意水準αはなぜ5%にすることが多いのか？

A. Fisherが言ったから。

R.A. Fisher

この「5%」が一人歩きすることは大いに問題

・5%を切らないと論文が採択されない？

・5%を切るまでサンプリングし、
　切ったとたんにサンプリングをストップして有意差を主張する？

アメリカ統計学会が警鐘を鳴らす論文を発表

R. L. Wasserstein and N. A. Lazar, *Am. Stat.* 70 (2016) 129.

「堂々巡り」“a long-worrisome circularity”
Q: なぜ大学では$p = 0.05$を教えるの?
A: 多くの人が使うから
Q: なぜ多くの人が$p = 0.05$を使うの？
A: 大学で教えているから

以下も参照

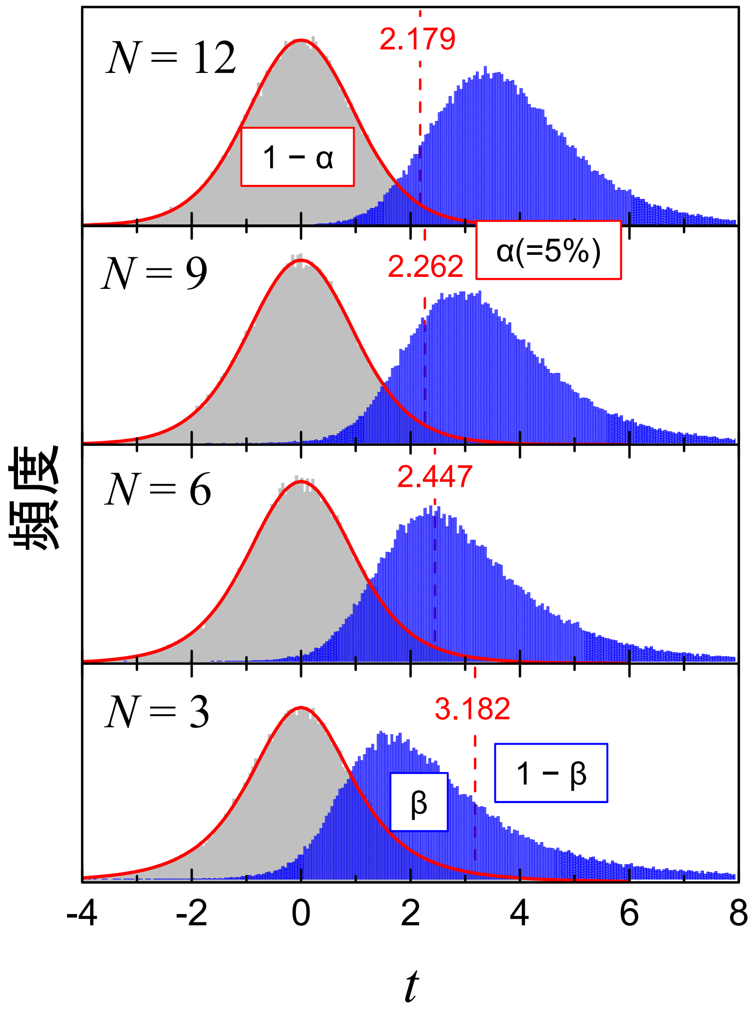
M. Baker, *Nature*, 531 (2016) 151.

諏訪雅頼, *ぶんせき*, 3(2017)120.

βの測定回数依存性

- ➢ $\mu_1 = \mu_0$（灰色）　　H_0が真
正規分布（$\mu_1 = 12.15$, $\sigma_0 = 0.05$）
からN回抽出、t値を計算（μ_0=12.15）

- ➢ $\mu_1 \neq \mu_0$（青色）　　H_1が真
正規分布（$\mu_1 = 12.20$, $\sigma_0 = 0.05$）
からN回抽出、t値を計算（μ_0=12.15）



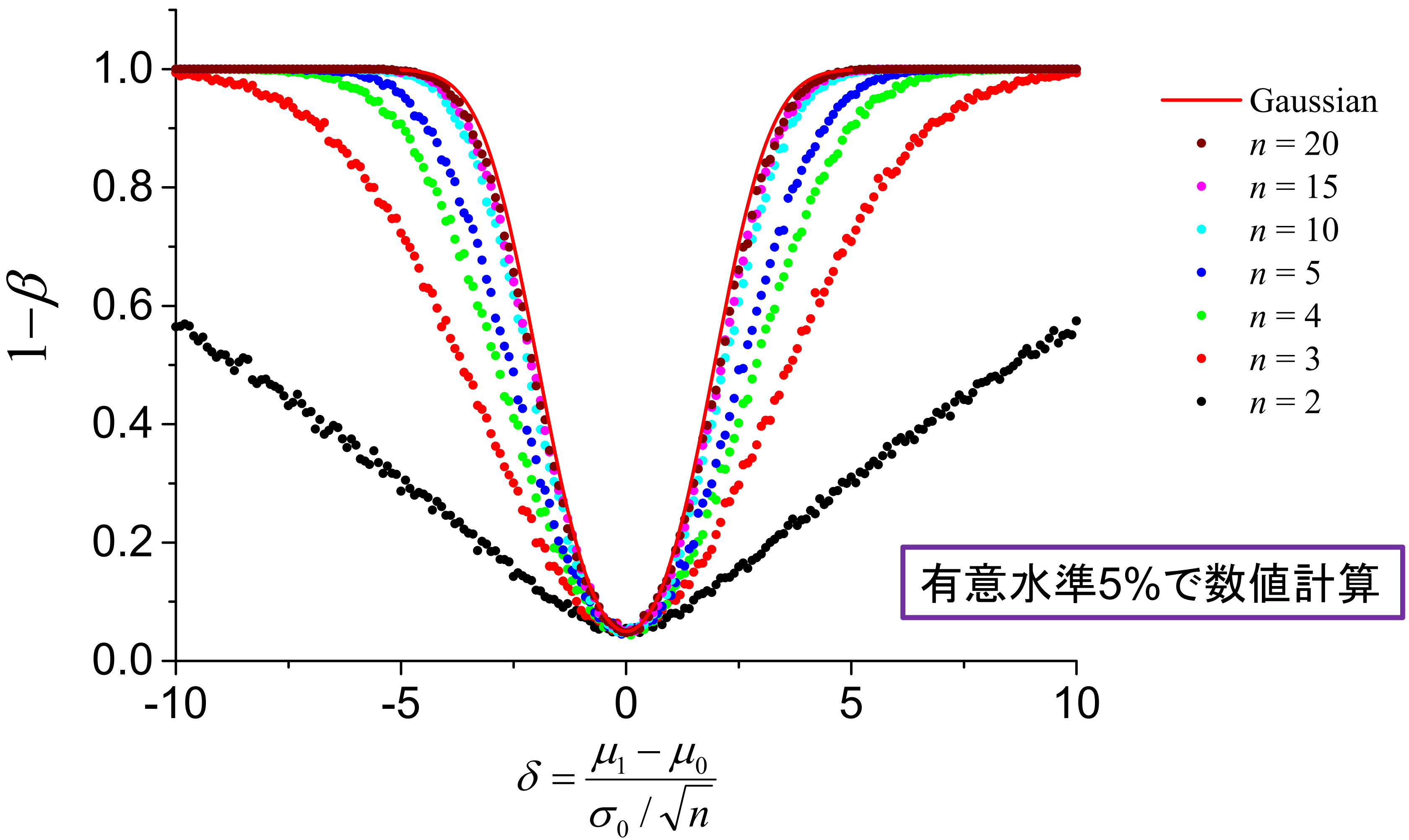
測定回数Nが多くなると
βは小さくなる

- 分散σ_0^2/Nが小さくなり、tの0からのずれが大きくなるため
- 一方で、臨界値の絶対値が小さくなるため

		検定結果	
		H_0を棄却できない	H_0を棄却しH_1を採用
事実	H_0が真	正しい 1 − α	第一種の誤り α（あわて）
	H_1が真	第二種の誤りβ（ぼんやり）	正しい 1 − β

検定力 $(1\text{-}\beta)$

有意差が検出される確率（検定力）



t検定II: 二つの測定値の比較

二つの測定値に有意差はあるか？

（母分散は未知）

【問題2-2】
下表は、同じ溶液中のある化学種の濃度を、A研究室とB研究室がそれぞれ独自に測定した結果である。

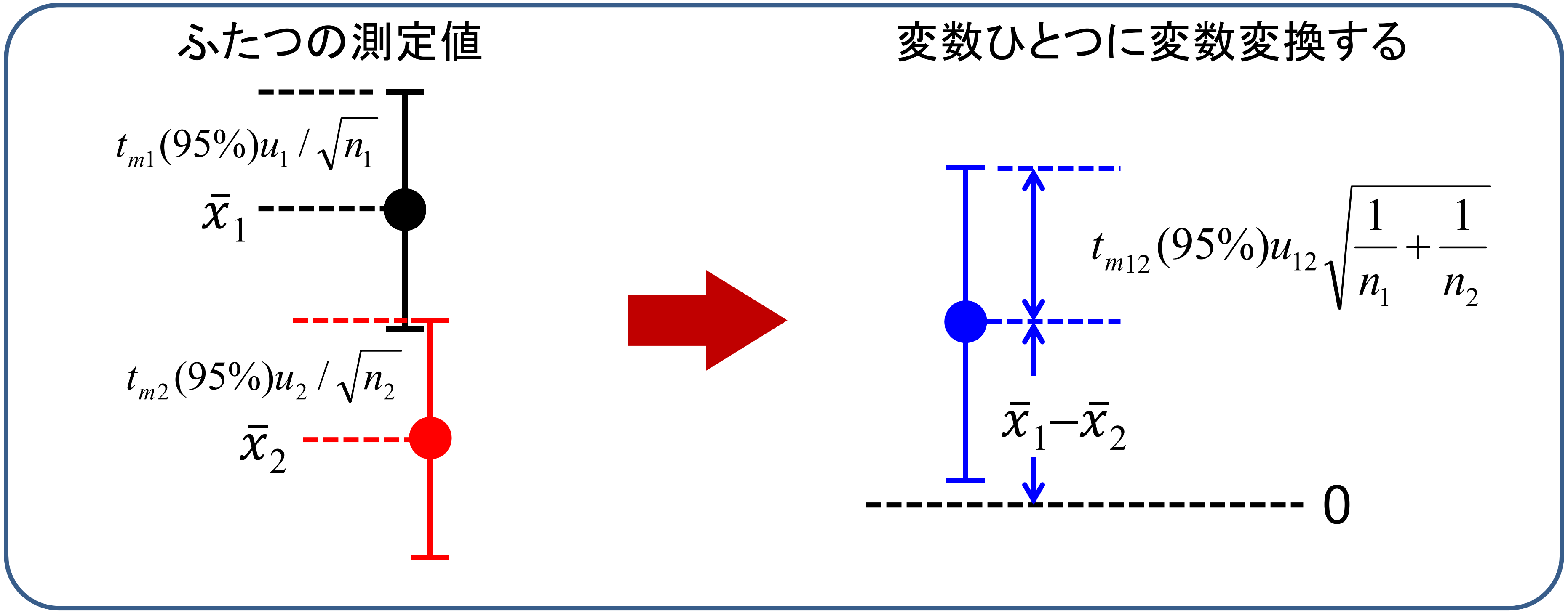
	平均 (ppm)	標準偏差 (ppm)	試行回数
A研究室	$\bar{x}_1 = 7.85$	$u_1 = 0.61$	5
B研究室	$\bar{x}_2 = 6.34$	$u_2 = 0.83$	7

測定値に有意差があるかどうかを有意水準5%で検定せよ。

1．適切な検定を決める

測定精度は未知 → 不偏分散から母分散を推定するt分布 → t検定

t検定II：平均値の加減算



もし$\sigma_1 = \sigma_2$であれば
合併不偏分散u_{12}が母分散の最良推定値

$$u_{12}^2 = \frac{(n_1-1)u_1^2 + (n_2-1)u_2^2}{n_1+n_2-2}$$

$$\boxed{t = \frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{u_{12}\sqrt{\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}}$$

自由度$m = n_1+n_2-2$のt分布に従う

t検定II：帰無仮説

流れ2．対立仮説／帰無仮説を決める

（仮定）$\sigma_1 = \sigma_2$

対立仮説H_1　$\mu_1 \neq \mu_2$
帰無仮説H_0　$\mu_1 = \mu_2$

両側検定

$$t = \frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{u_{12}\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \boxed{\frac{\bar{x}_1 - \bar{x}_2}{u_{12}\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}}$$

H_0が正しいなら、自由度$m = n_1 + n_2 - 2$のt分布に従う

先立って$\sigma_1 = \sigma_2$であることを確認する必要がある　⇒　F検定（後述）

t検定II：手順

流れ3．有意水準 α を決め、臨界値を計算する

自由度 $m = n_1 + n_2 - 2 = 10$

$t_{10,5\%} = \pm 2.228$

”T.INV.2T(0.05,10)”

流れ4．帰無仮説を仮定して統計量Xを計算する

$$u_{12} = \sqrt{\frac{(n_1-1)u_1^{\ 2} + (n_2-1)u_2^{\ 2}}{n_1+n_2-2}} = \sqrt{\frac{(3-1)\times 0.61^2 + (5-1)\times 0.83^2}{5+7-2}} = 0.750$$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{u_{12}\sqrt{\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{7.85-6.34}{0.750\times\sqrt{\frac{1}{5}+\frac{1}{7}}} = 3.44$$

流れ5．判定

3.44 > +2.228 …有意水準5%で有意差あり

χ^2検定：分散の基準値に対する比較

測定値の分散は基準値に対して有意差があるか？

【問題0-3】

工場のラインにおいて、製品の品質を反映する物理量 x を自動測定して、品質管理をしている。x の自動測定結果は、母分散 $\sigma_0^2 = 10^2$の正規分布に従うことがわかっている。正常な製品の x は母平均 $\mu_0 = 50$を示す。

x のずれをより良く検出するために自動測定を改良した。改良後に、製品10個で不偏分散 u^2 を評価したところ $u^2 = 5.0^2$であった。測定精度は上がったたと言えるだろうか？　有意水準5%で判断せよ。

流れ1．適切な検定を決める

不偏分散と分散基準値の比較 → χ^2分布を用いる → χ^2検定

χ^2分布

$x_1, x_2, \ldots, x_m$が独立に正規分布に従うとき

$$\boxed{\sum_{i=1}^{m}\left(\frac{x_i-\mu}{\sigma}\right)^2 = \sum_{i=1}^{m} z_i^{\,2} \equiv \chi_m^2}$$

自由度mのχ^2分布
これは定義

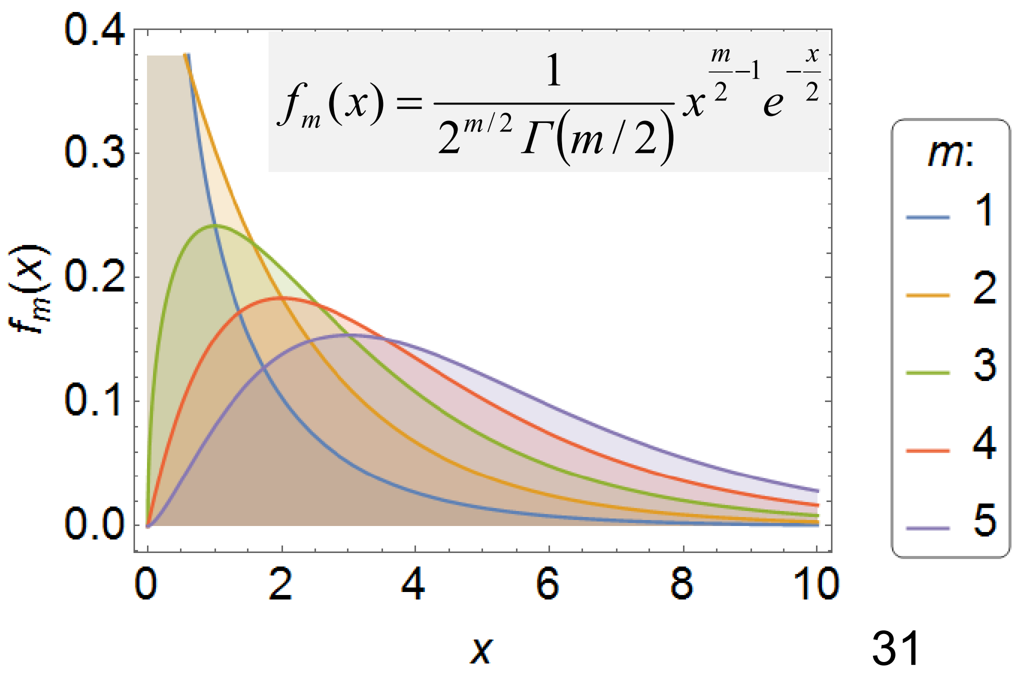
$$\boxed{\sum_{i=1}^{m}\left(\frac{x_i-\bar{x}}{\sigma}\right)^2 = \left(\frac{m-1}{\sigma^2}\right)u^2 = \chi_{m-1}^2}$$

自由度$m-1$のχ^2分布
χ^2分布は不偏分散u^2の分布

- χ^2分布はばらつきの指標となる

- $\mathrm{E}[\chi_m^{\ 2}] = m$
  $\chi_m^{\ 2}$はだいたいmになる

H_0が真ならば、計算したχ^2は
mあたりに値を持つはず



χ^2検定：対立仮説/帰無仮説、臨界値

流れ2．対立仮説／帰無仮説を決める

対立仮説 H_1:　$\mathrm{E}[u^2] = \sigma_1^2 < \sigma_0^2$

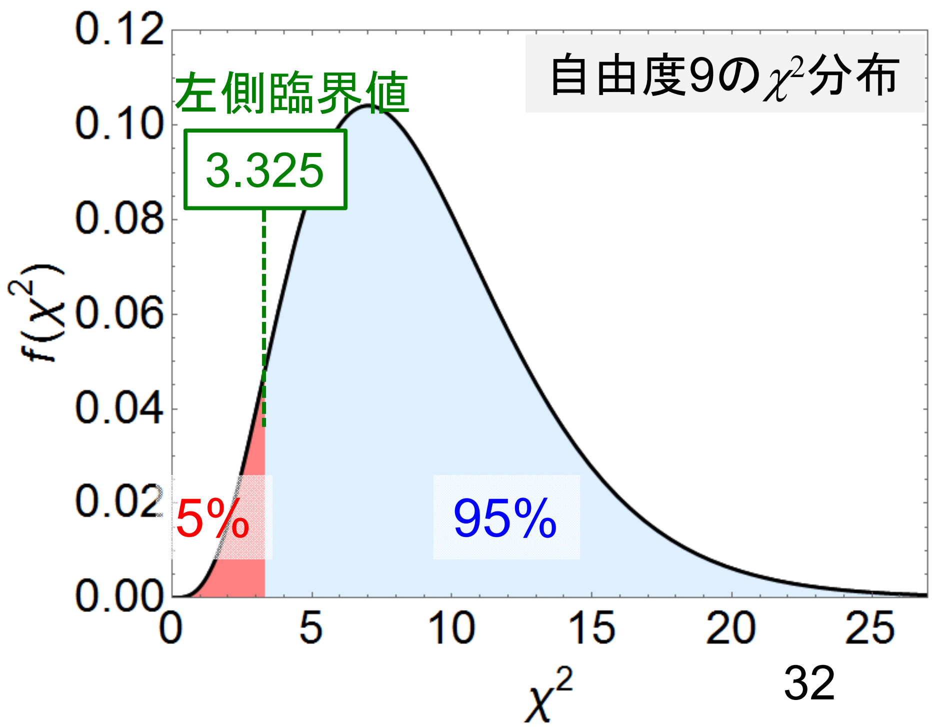
帰無仮説 H_0:　$\mathrm{E}[u^2] = \sigma_1^2 = \sigma_0^2$

u^2が小さいかどうかを検定したいので、片側検定（左側）

流れ3．有意水準 α を決め、臨界値 $\chi^2_{N-1,\alpha}$を計算する

$$\chi^2_{10-1,5\%,\mathrm{left}} = 3.325$$

"CHISQ.INV(0.05,9)"



χ^2検定：判定

流れ4．帰無仮説を仮定して統計量Xを計算する

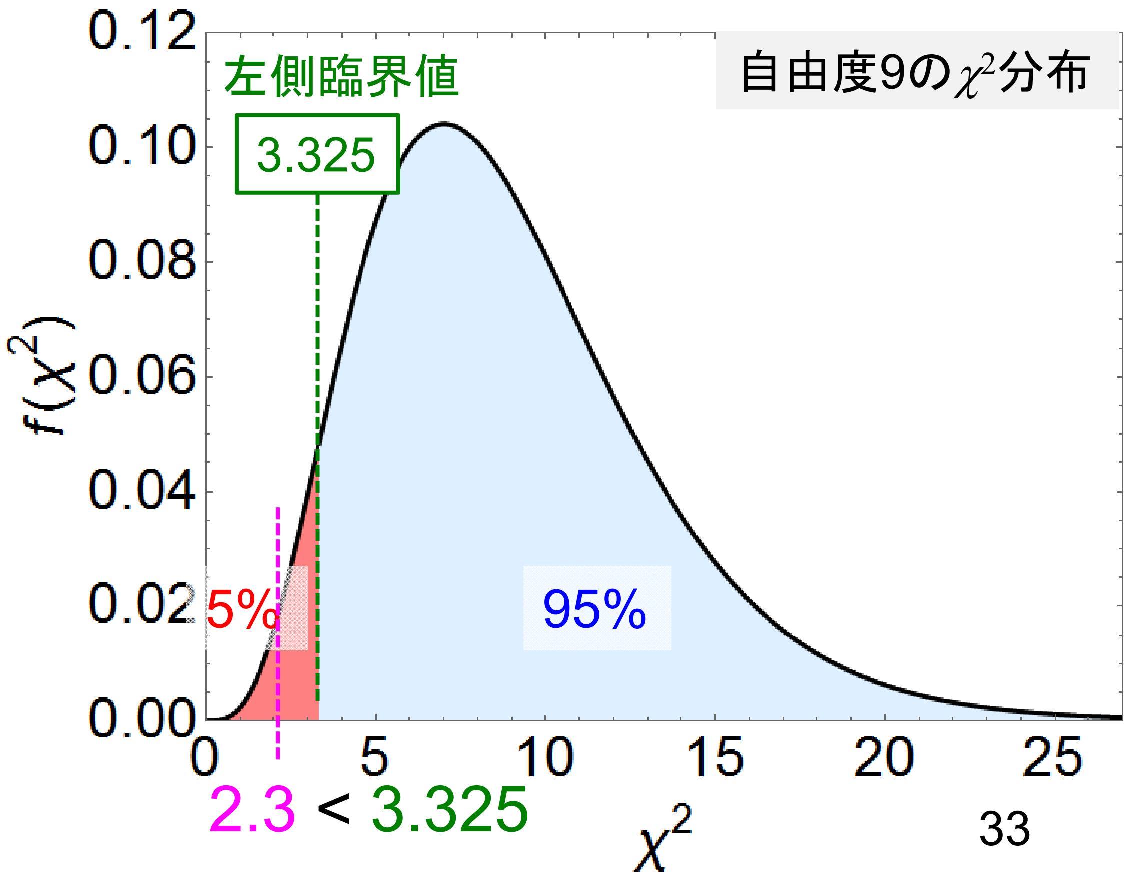
測定回数 N=10
有意水準　α=5%
信頼度　1-α=95%

$$\chi^2 = \frac{N-1}{\sigma_0^2}u^2 = \frac{10-1}{10^2}5.0^2 = 2.3$$

流れ5．判定

2.3 < 3.325
よりH_0を棄却し、H_1を採用する

自動測定の改良により測定精度が上がったと言える



F検定：二つの測定値の分散の比較

二つのデータ群の分散に有意差はあるか？

【問題2-1】
下表は、同じ溶液中のある化学種の濃度を、A研究室とB研究室がそれぞれ独自に測定した結果である。

	平均 (ppm)	標準偏差 (ppm)	測定回数
A研究室	$\bar{x}_1$ = 7.85	u_1 = 0.61	5
B研究室	$\bar{x}_2$ = 6.34	u_2 = 0.83	7

測定精度に有意差があるかどうかを有意水準5%で検定せよ。

流れ1．適切な検定を決める

ふたつの不偏分散の比較 → F分布を用いる → F検定

F分布

$$F=\frac{u_1^{\ 2}/\sigma_1^{\ 2}}{u_2^{\ 2}/\sigma_2^{\ 2}}=\frac{\chi_1^{\ 2}/m_1}{\chi_2^{\ 2}/m_2}$$

F分布の定義:カイ二乗分布に従う2つの変数の比

検定で使う形:母分散で規格化した不偏分散の比

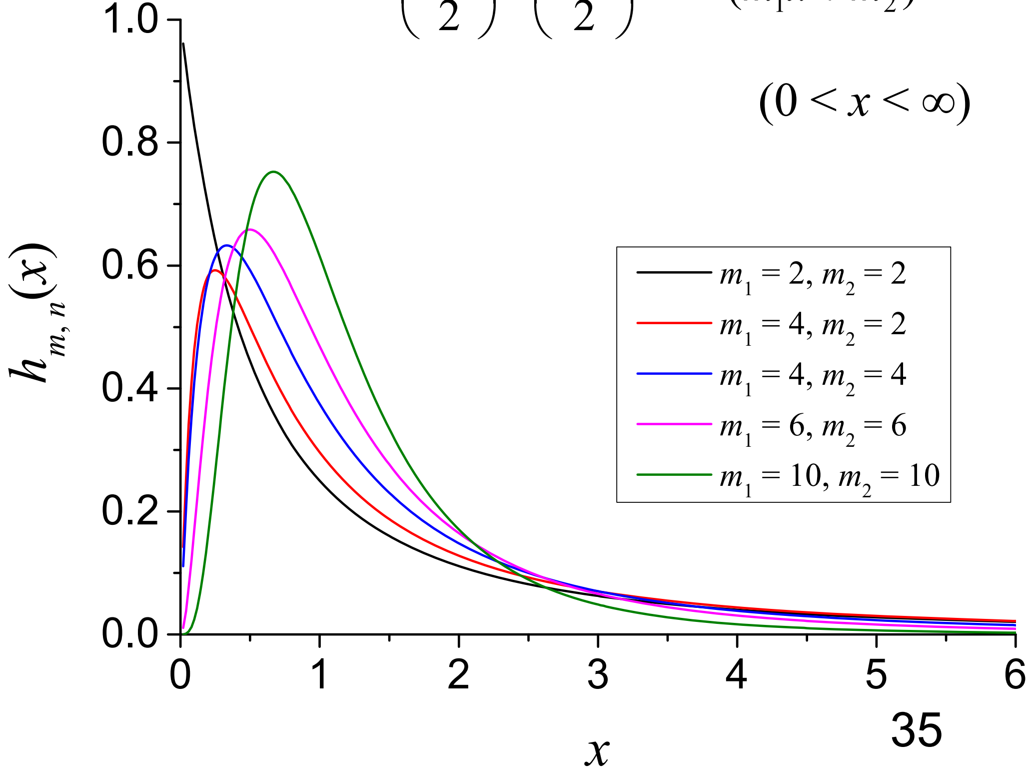
・F分布は二つのばらつきの比

・F値の中央値はだいたい1

H_0が真ならば、計算したFは1から遠くない値を持つはず

$$h_{m_1,m_2}(x)=\frac{\Gamma\left(\frac{m_1+m_2}{2}\right)m_1^{\frac{m_1}{2}}m_2^{\frac{m_2}{2}}}{\Gamma\left(\frac{m_1}{2}\right)\Gamma\left(\frac{m_2}{2}\right)}\cdot\frac{x^{\frac{m_1}{2}-1}}{(m_1x+m_2)^{\frac{m_1+m_2}{2}}}$$

$$(0<x<\infty)$$



F検定: 帰無仮説

流れ2．対立仮説／帰無仮説を決める

対立仮説 H_1:　$\sigma_1^2 \neq \sigma_2^2$
帰無仮説 H_0:　$\sigma_1^2 = \sigma_2^2$

有意差があるかないかを検定したいので、両側検定

H_0が成り立つとき

$$\boxed{F = \frac{u_1^2 / \sigma_1^2}{u_2^2 / \sigma_2^2} = \frac{u_1^2}{u_2^2}}$$

分子の自由度:　m_1 = 7 - 1 = 6
分母の自由度:　m_2 = 5 - 1 = 4

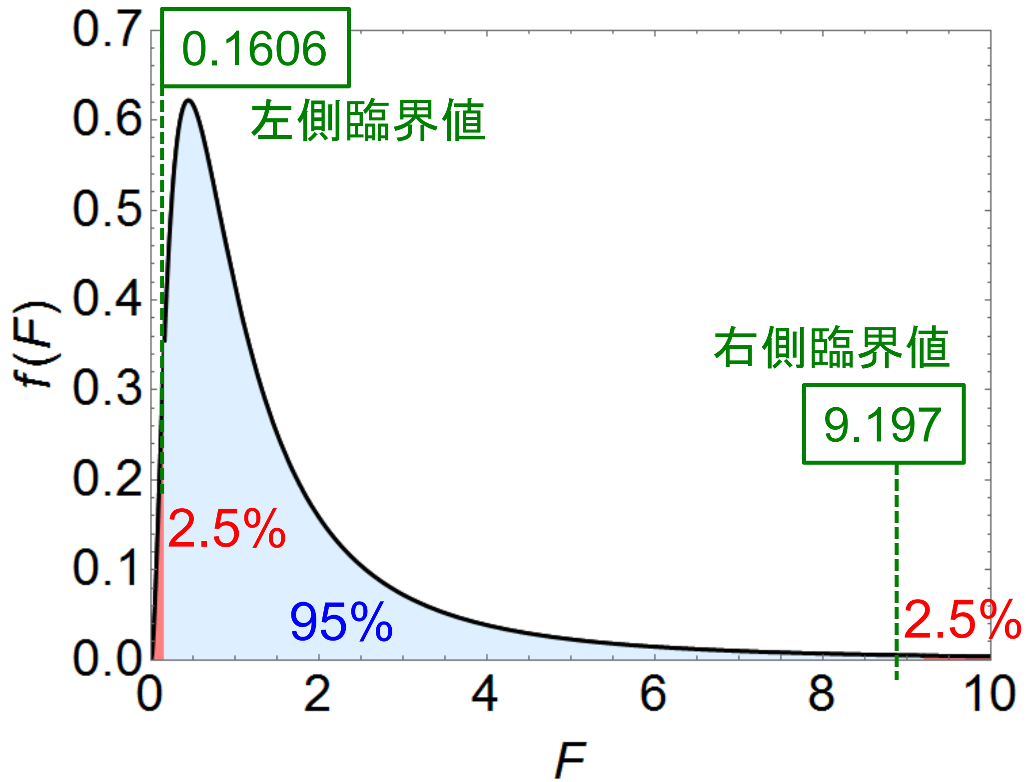
F検定: 有意水準と臨界値

流れ3．有意水準 α を決め、臨界値$F_{m,n,\alpha}$ を計算する

分子の自由度： m_1 = 7 - 1 = 6
分母の自由度： m_2 = 5 - 1 = 4

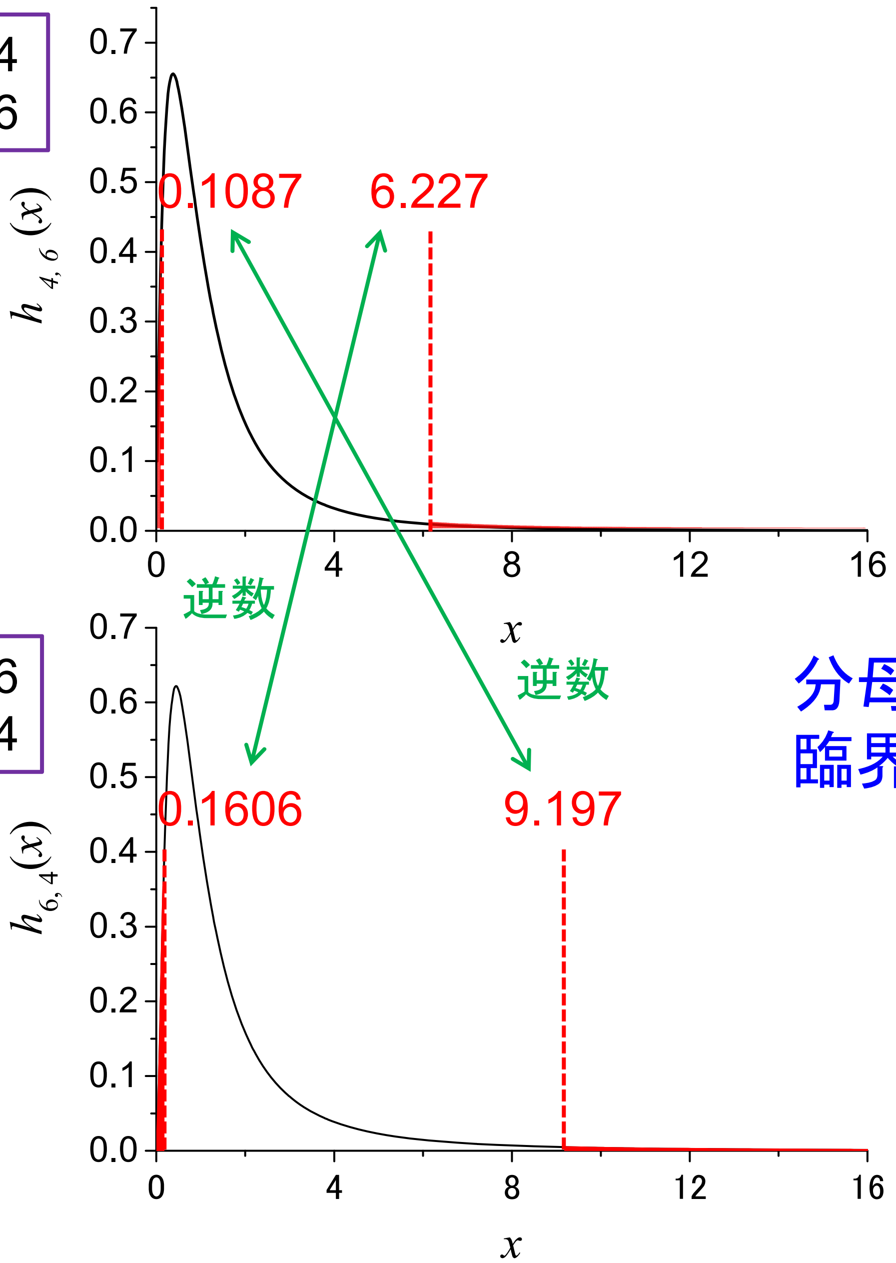
$$F_{7-1,5-1,5\%} = 0.1606,\ 9.197$$

左側臨界値 “F.INV(0.025,6,4)”
右側臨界値 “F.INV.RT(0.025,6,4)”



F検定：分母と分子

分子の自由度：4
分母の自由度：6



分子の自由度：6
分母の自由度：4

分母分子を逆にすると
臨界値が変わる

F検定：手順

流れ3．有意水準 α を決め、臨界値を計算する

分子の自由度 m_2 = 7 – 1 = 6
分母の自由度 m_1 = 5 – 1 = 4

$F_{6,4,5\%}$ = 0.1606, 9.197

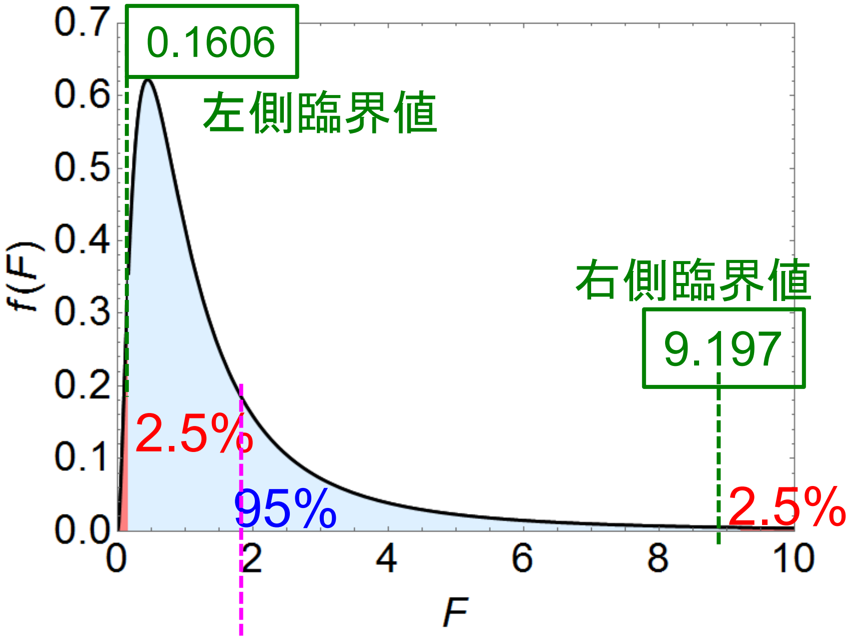
"F.INV(0.025,6,4)", "F.INV.RT(0.025,6,4)"

流れ4．帰無仮説を仮定して
統計量を計算する

$$F = \frac{u_2^{\ 2}}{u_1^{\ 2}} = \frac{0.83^2}{0.61^2} = 1.85$$

流れ5．判定

0.1606 < 1.85 < 9.197…有意差なし



参考図書（緑本以外）

1. ゼロから学ぶ統計解析：小寺

2. 推定と検定のはなし：蓑谷

3. 実践統計学入門：足立

4. 統計学が最強の学問である［実践編］：西内

5. 最小二乗法の理論とその応用（改訂版）：田島、小牧

6. Introduction to Statistical Analysis: Dixon, Massey

最小二乗法と検量線

大阪大学　大学院理学研究科
諏訪雅頼
msuwa@chem.sci.osaka-u.ac.jp

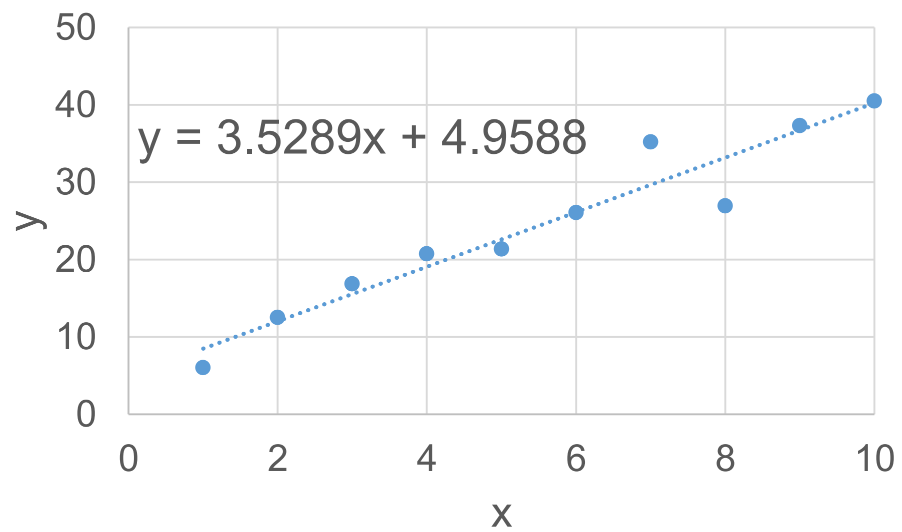
内容

- 最小二乗法に必要な基礎知識
  偶然誤差、正規分布、残差、残差の二乗和、χ^2
- 線形最小二乗法（重みなし）
- 逆推定の誤差
- 当てはまりの良さ、モデルのもっともらしさ
- 線形最小二乗法（重み付き）

回帰分析でよくある例

エクセルでグラフを描き，近似直線を表示，

パラメタをコピーするが，



× 桁が異常に多い

$y = 3.5289x + 4.9588$

△ なんとなく二桁に丸める

$y = 3.5x + 5.0$

○ 最小二乗法でパラメタの誤差まで検討

a=3.5±0.4, b=5±2 　($y = ax + b$)

◎ モデルのもっともらしさを検証

AICcの値より、$y=ax$でも$y=ax^2+bx+c$でもなく、$y=ax+b$が最ももっともらしいモデル

モデル	AICc
$y=ax$	60.1
$y=ax+b$	**59.5**
$y=ax^2+bx+c$	65.0

最小二乗法とは？

- 最小二乗法とは？

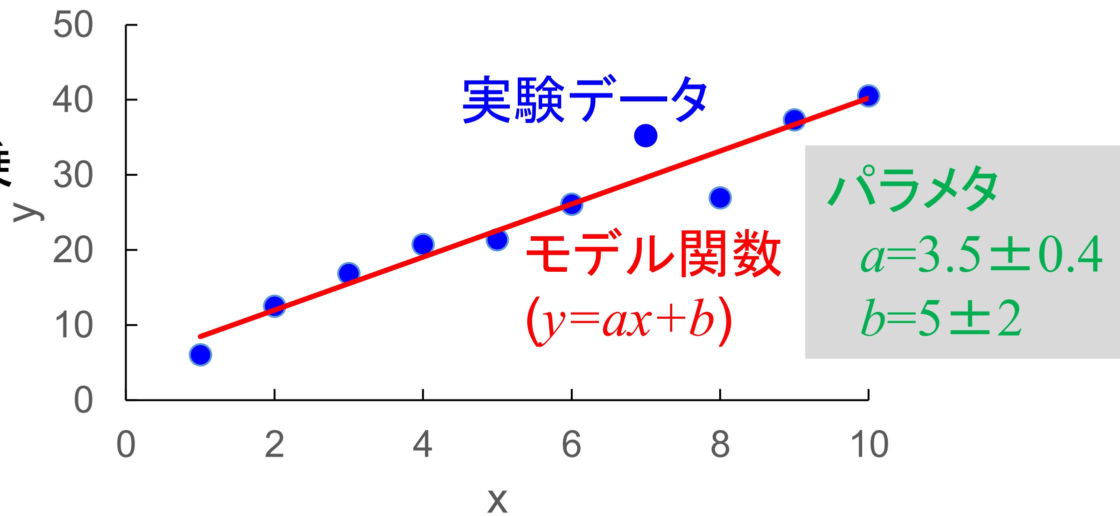
  実験データ （e.g. $\{(x_1, y_1),\ldots(x_n, y_n)\}$） に
  モデル関数 （e.g. $y=ax+b$） をフィットさせて
  パラメタ （e.g. a,b） の最良推定値を決定する方法

- 何が二乗？

  残差（ $\varepsilon_i = y_i-(ax_i+b)$ ）が二乗

- 何が最小？

  残差の二乗和（$\Sigma\varepsilon_i^2$）が最小



- なぜ？

  その実験データが測定される確率が最も高いモデルなので

  最尤原理「手元に得られた標本は、最も得やすい標本である」

本講習

- 本講習：主に、$y=ax+b$による線形最小二乗法・重みなしをやります

  最小二乗法の基礎中の基礎。基礎ができれば応用もできる

  エクセルに関数あり（LINEST）。しかし、自分で一度は手計算しよう

- 線形or非線形

  線形とはパラメタの線形結合という意味。$y=ax+b$の１次関数の線形ではない。

  線形　の例：$y=ax+b$, $y=ax^2+bx+c$, $y=a\cos x+b\sin x$

  非線形の例：$y=a\exp(-bx)+c$, $y=a\sin(bx+c)$

- 重みなしor重み付き

  重みなし：データのばらつき具合がデータによらず一定

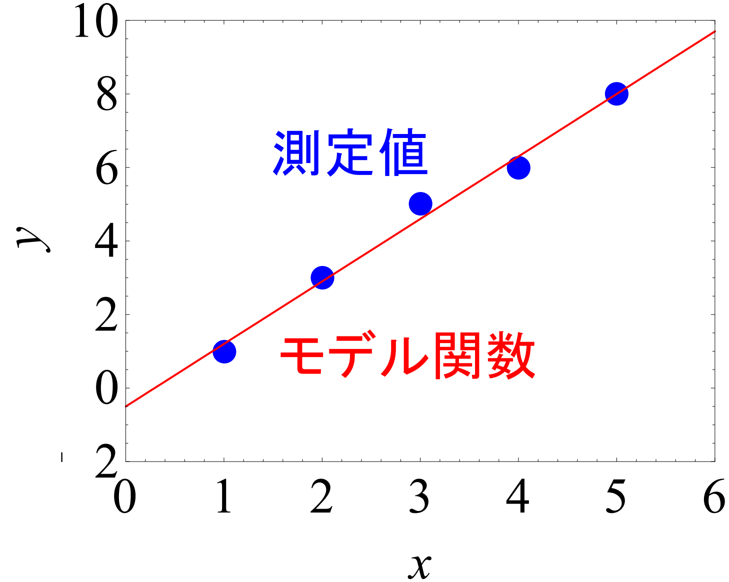
  重み付き：データのばらつき具合がデータに依存し、“重み”がかかる

残差　ε_i

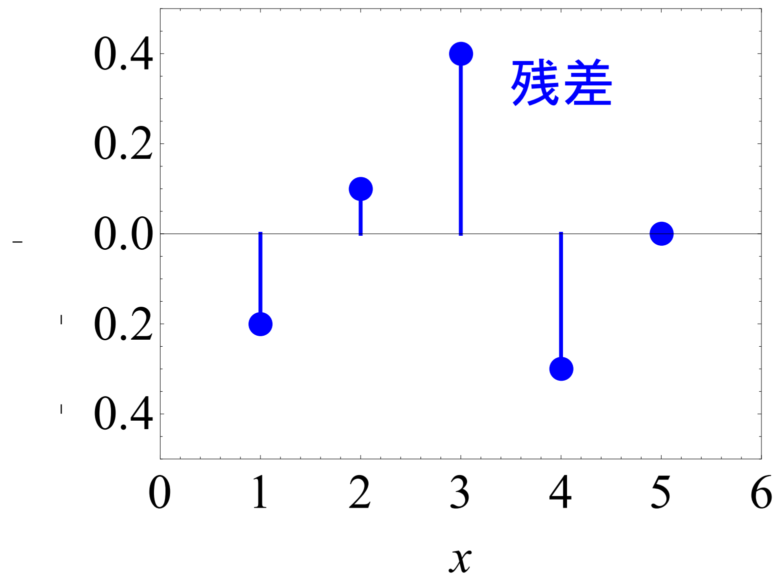
- 残差ε_i
  測定値y_iのモデル関数$f(x_i)$からのずれ

$$\varepsilon_i = y_i - f(x_i)$$

e.g. $f(x_i) = ax_i + b$



- 最小二乗法における仮定
  - 各データ測定は独立試行であり、
    各データの残差はそれぞれ独立
  - 残差は偶然誤差からなる
    （系統誤差を含まない）
  - 残差の母集団は正規分布をなす



ランダム誤差の評価：正規分布

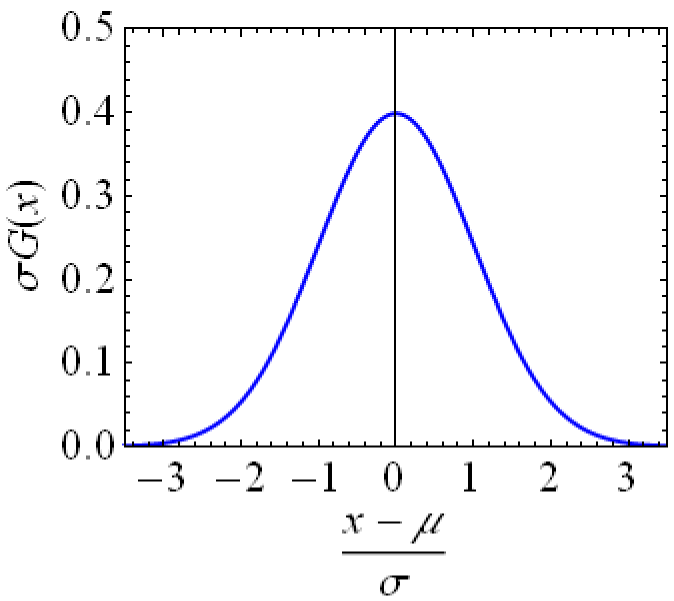
➢ランダム誤差が正規分布で表わされると仮定

$$G_{\mu,\sigma}(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$$

μ：平均

σ^2：分散

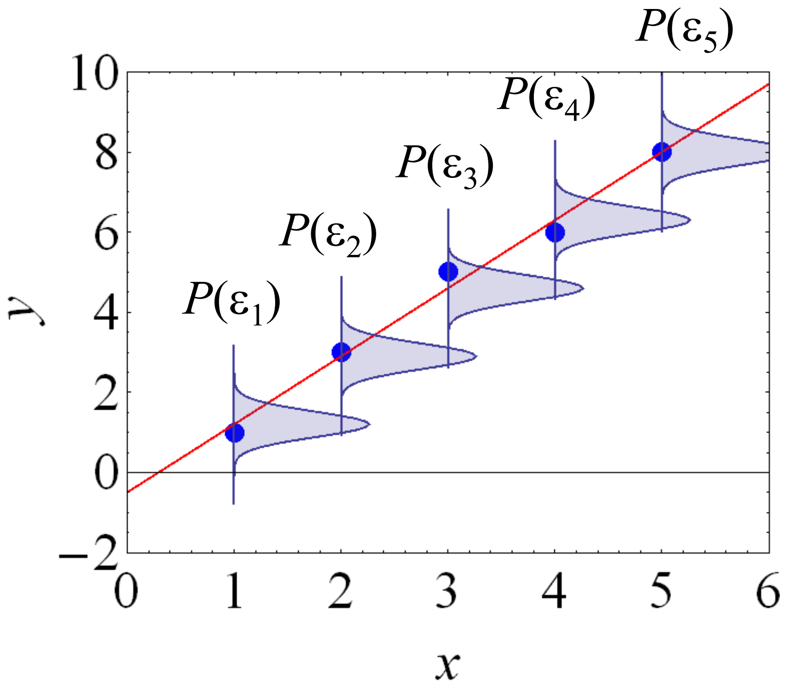
σ：標準偏差

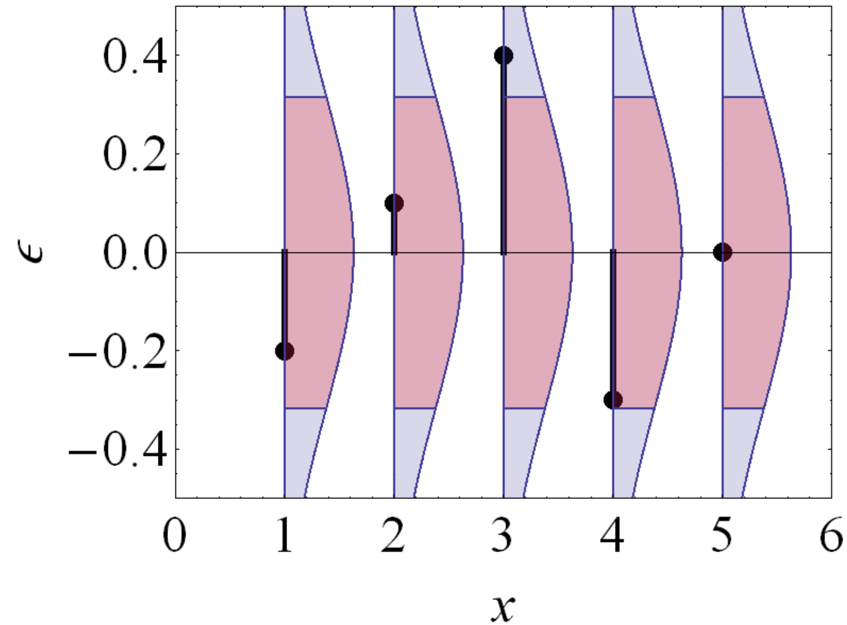


手元の測定値が得られる確率　$P(\varepsilon_i)$

- 測定の分散をσ_{yi}^2とすると、測定値の残差がε_iになる確率$P(\varepsilon_i)$は、

$$P(\varepsilon_i) \propto \frac{1}{\sigma_{yi}} \exp\left[-\frac{\varepsilon_i^2}{2\sigma_{yi}^2} \right]$$





手元の測定データが得られる確率　P_{total}

- 測定データの組$\{(x_1, y_1),\ldots,(x_n, y_n)\}$が測定される確率$P_{\text{total}}$は$P(\varepsilon_i)$の積

$$P_{total} = \prod_{i=1}^{n} P(\varepsilon_i)$$

$$\propto \frac{1}{\prod \sigma_{yi}} \exp\left[-\frac{1}{2} \sum_{i=1}^{n} \frac{\varepsilon_i{}^2}{\sigma_{yi}{}^2} \right]$$

$$= \frac{1}{\prod \sigma_{yi}} \exp\left[-\frac{1}{2} \chi^2 \right]$$

これが最大になるようなパラメタ値が、最ももっともらしい（最尤）

$$\chi^2 = \sum_i \frac{\varepsilon_i^2}{\sigma_{yi}^2}$$

分散で規格化された残差二乗和

= 標準正規変数の二乗和

正規方程式（重みなし）

- 最尤 → P_{total} 最大 → χ^2最小

  χ^2 最小のときの$\{a,b\}$を求める（ここでは重みなし$\sigma_{yi}=\sigma_y$を仮定）

$$\frac{\partial\chi^2}{\partial a}=\frac{-2}{{\sigma_y}^2}\sum_{i=1}^{N}x_i\{y_i-(ax_i+b)\}=0$$

$$\frac{\partial\chi^2}{\partial b}=\frac{-2}{{\sigma_y}^2}\sum_{i=1}^{N}\{y_i-(ax_i+b)\}=0$$

この$\{a,b\}$の二元連立一次方程式を解いて$\{a,b\}$の最良推定値を求める

χ^2を最小にするm元連立一次方程式を正規方程式という。

（パラメタの個数mが多い場合、手計算は大変。行列をつかう。）

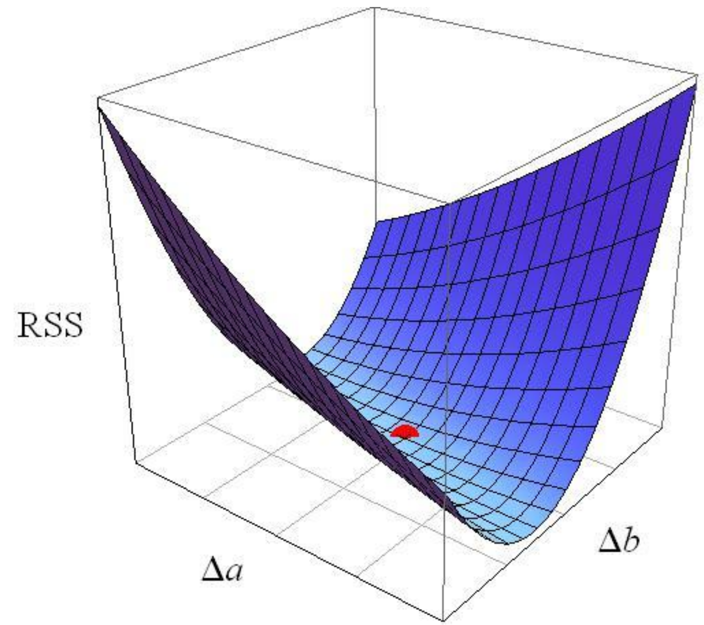
パラメタの最良推定値

➢正規方程式を解くと

$$a = \frac{N\Sigma xy - \Sigma x \Sigma y}{\Delta}$$

$$b = \frac{\Sigma x^2 \Sigma y - \Sigma x \Sigma xy}{\Delta}$$

$$\Delta = N\Sigma x^2 - (\Sigma x)^2$$



（ $\sum_{i=1}^{N} x_i$ を $\sum x$ と略記している）

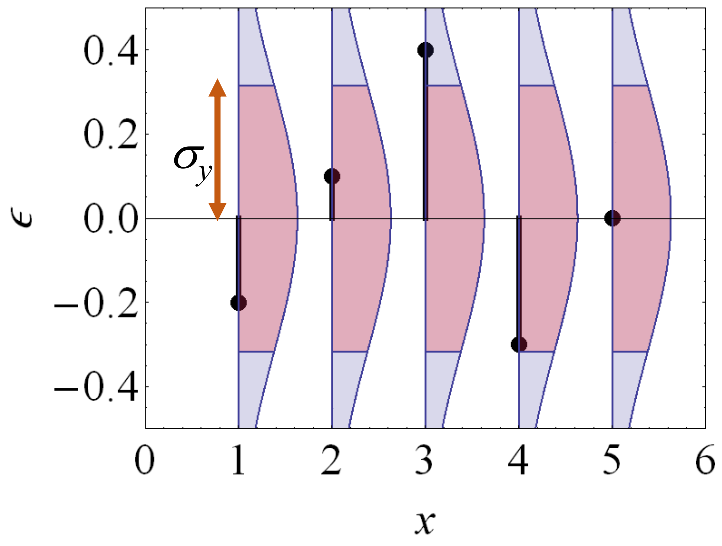
これが$\{a,b\}$の最良推定値であり、ここから$\{a,b\}$の値をどのようにずらしても残差の二乗和（RSS）は上昇する。

パラメタの誤差の推定

- 測定値yの不偏標準偏差

$$u_y = \sqrt{\frac{\sum_i {\varepsilon_i}^2}{N-2}}$$

⇐ yの母標準偏差σ_yの最良推定値



- 傾きaと切片bの不偏標準偏差

$$u_a = \sqrt{\sum_i \left(\frac{\partial a}{\partial y_i}\right)^2 (u_y)^2} = u_y \sqrt{\frac{N}{\Delta}}$$

$$u_b = \sqrt{\sum_i \left(\frac{\partial b}{\partial y_i}\right)^2 (u_y)^2} = u_y \sqrt{\frac{\sum x^2}{\Delta}}$$

- 傾きと切片の信頼区間

$$(a-\mu_a)/u_a \,, (b-\mu_b)/u_b$$

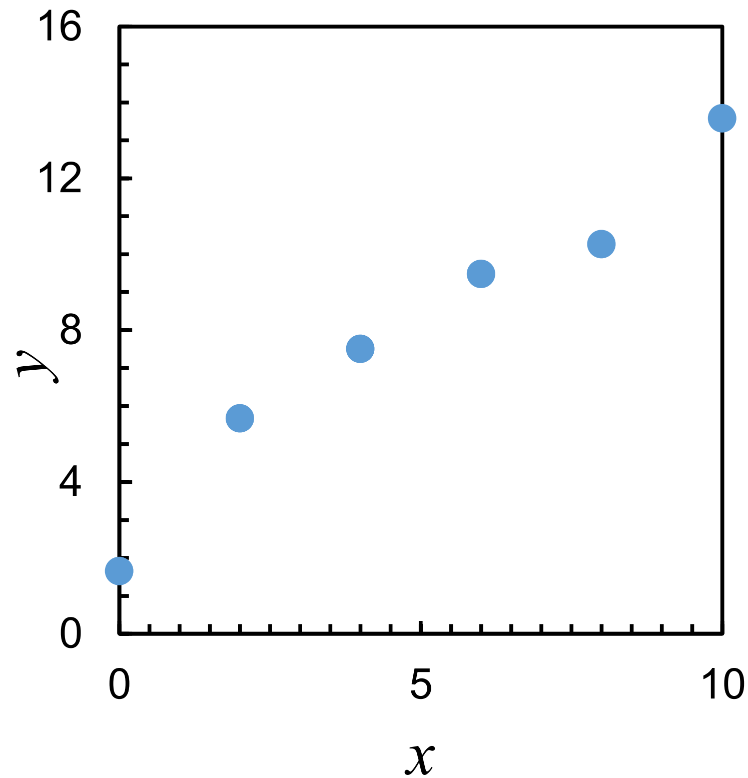
⇒自由度N–2のt分布に従う

$$\delta a = t_{N-2,\alpha} u_a$$
$$\delta b = t_{N-2,\alpha} u_b$$

例題　線形最小二乗法・重み無し

ある信号の強度yと試料濃度xの関係を調べると、下表の結果が得られた。測定値に直線を当てはめて傾きと切片を見積もる。

x_i	y_i
0.00	1.64
2.00	5.67
4.00	7.50
6.00	9.47
8.00	10.26
10.00	13.58

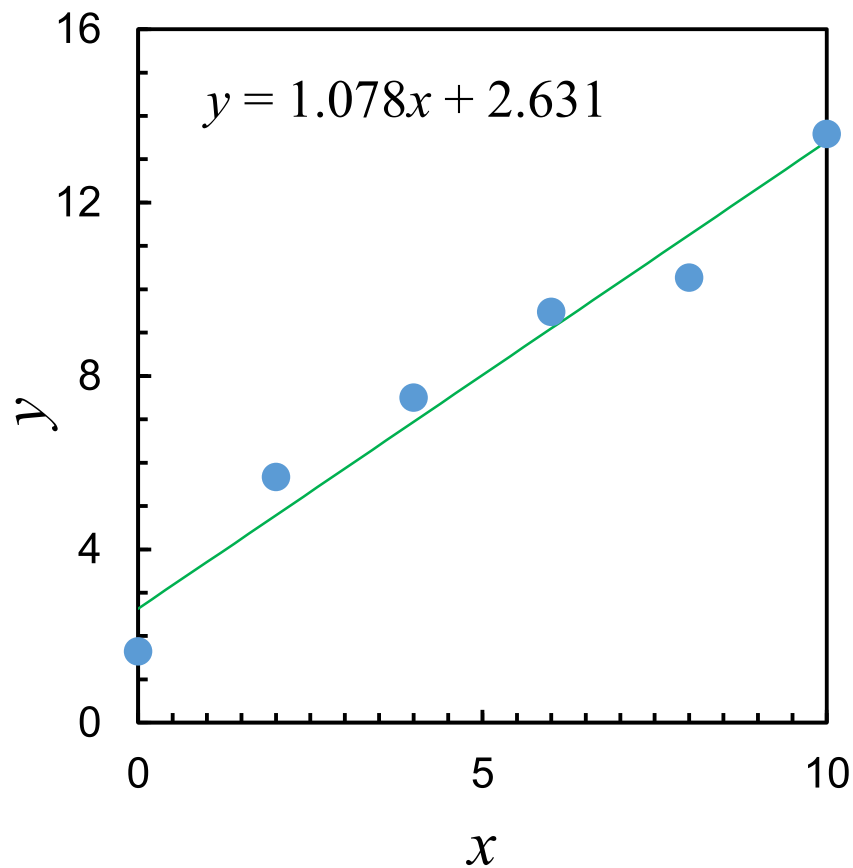


エクセル関数を使わずに最小二乗法を行ってみましょう

例題　線形最小二乗法・重み無し　パラメタ算出

1. 測定データからΣx, Σy, Σx^2, Σxyを計算

	x_i	y_i	x_i^2	$x_i y_i$
	0.00	1.64	0.00	0.00
	2.00	5.67	4.00	11.34
	4.00	7.50	16.00	30.00
	6.00	9.47	36.00	56.82
	8.00	10.26	64.00	82.08
	10.00	13.58	100.00	135.80
Σ	30.00	48.12	220.00	316.04



2.　正規方程式の解に代入してa, bを計算

$$\Delta = 6 \times 220.00 - (30.00)^2 = 420.00$$

$$a = \frac{6 \times 316.04 - 30.00 \times 48.12}{420.00} = 1.078$$

$$b = \frac{220.00 \times 48.12 - 30.00 \times 316.04}{420.00} = 2.631$$

パラメタの有効数字は？

例題　線形最小二乗法・重み無し　パラメタの誤差

3. a, bからRSS$(\Sigma\varepsilon^2 = \Sigma\{y-(ax+b)\}^2)$を計算

	x_i	y_i	$ax_i + b$	ε_i^2
	0.00	1.64	2.632	0.980
	2.00	5.67	4.787	0.779
	4.00	7.50	6.943	0.310
	6.00	9.47	9.098	0.138
	8.00	10.26	11.254	0.988
	10.00	13.58	13.409	0.029
Σ	30.00	48.12		3.228

4. yの不偏標準偏差u_yを計算

$$u_y = \sqrt{\frac{3.228}{6-2}} = 0.8984$$

5. u_yからu_aとu_bを計算

$$u_a = 0.8983\sqrt{\frac{6}{420}} = 0.1077$$

$$u_b = 0.8983\sqrt{\frac{220}{420}} = 0.6502$$

6. それぞれの信頼区間を計算

信頼度α = 95%とすると,

$$\delta a = t_{4,95\%} u_a = 2.776 \times 0.1071 = 0.30$$

$$\delta b = t_{4,95\%} u_b = 2.776 \times 0.6484 = 1.8$$

誤差を一桁で丸めて

$y = (1.1 \pm 0.3)x + (3 \pm 2)$

ExcelのLINEST関数(配列関数)

LINEST(データy,データx, [定数], [補正])

$y = ax+b$ ([定数]がTRUE)または$y = ax$ ([定数]がFALSE)をモデル関数として線形最小二乗法フィッティング(重みなし)を行う関数,

x	y
0.000	0.0000
0.100	0.2517
0.200	0.4970
0.300	0.7553
0.400	1.0086

LINEST(yのデータ,xのデータ,TRUE,TRUE)

2.5208	−0.00164
0.011391	0.00279
0.999939	0.003602

=

a	b
u_a	u_b
R^2	u_y

[補正]がTRUEの場合に表示される

適当なセルに数式を入力

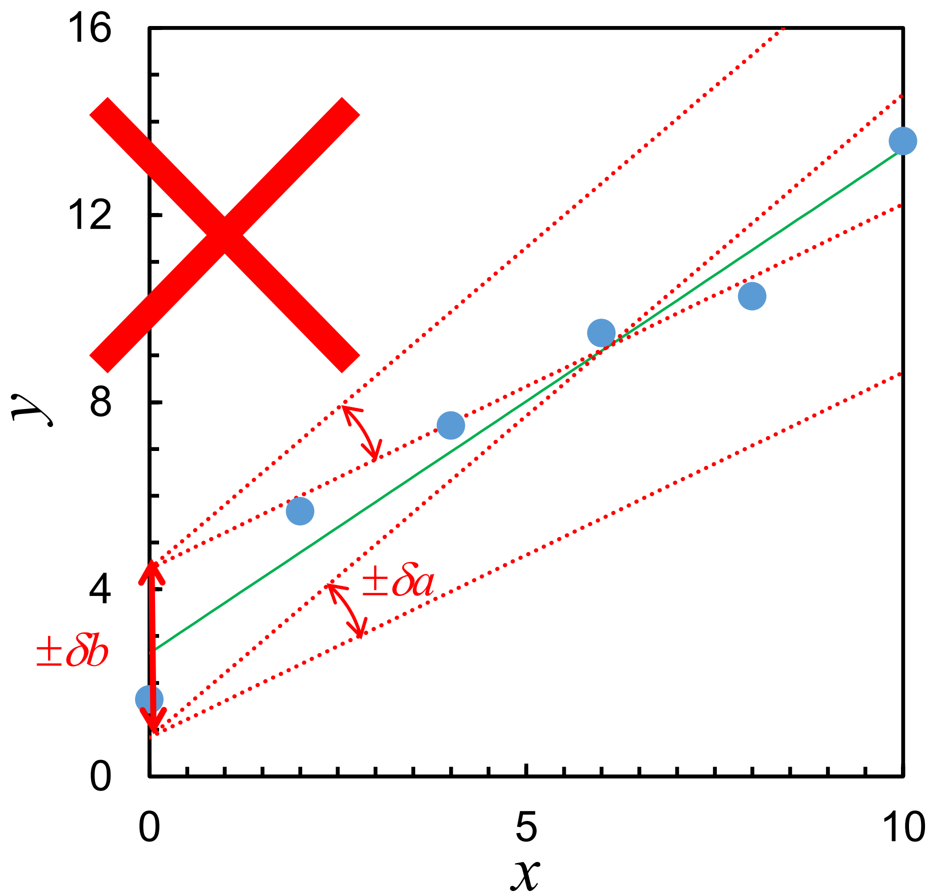
数式を左上として3行2列の出力範囲を選択

F2を押してからShift+Ctl+Enter

回帰直線の誤差を図示（誤り）

$y = (1.1 \pm 0.3)x + (3 \pm 2)$

切片を中心に傾きを変化させてみる



回帰直線の誤差を図示（正解）

傾き回転の中心はxの平均値$\bar{x}$

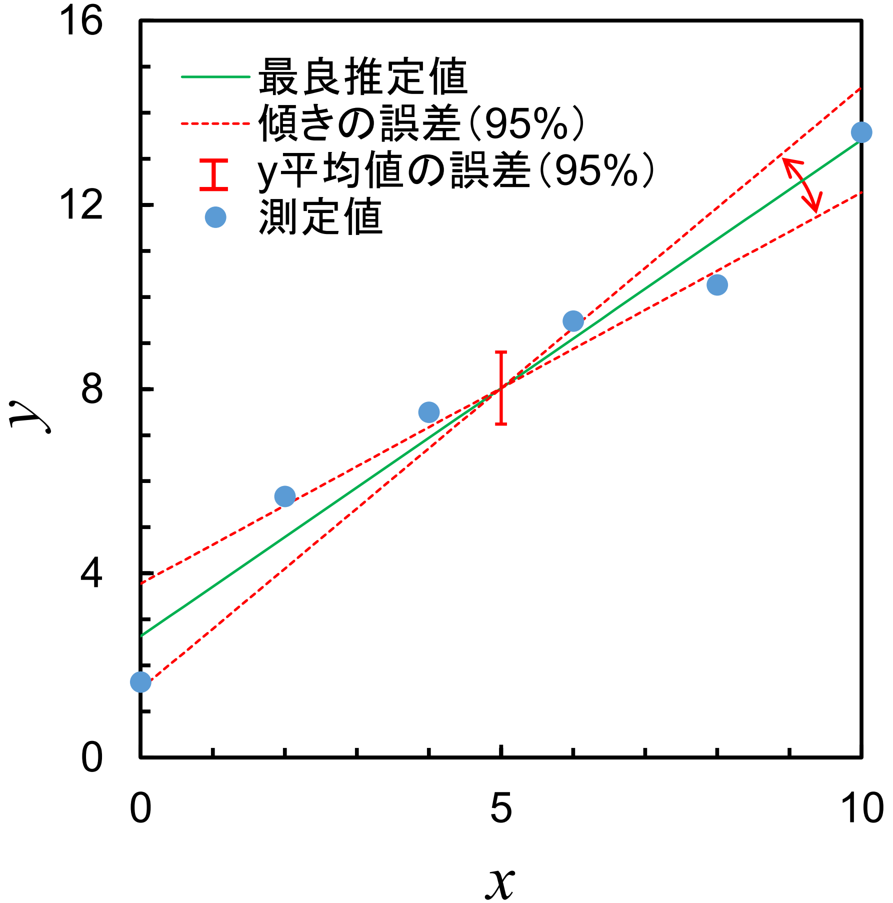
正規方程式（bで偏微分した方）から

$$b = \frac{\sum_i y_i}{N} - a\frac{\sum_i x_i}{N} = \bar{y} - a\bar{x}$$

bの代わりにy_iの平均値$\bar{y}$を使う

$$y = ax + b$$

$$\Rightarrow\ y - \bar{y} = a(x - \bar{x})$$



- $\bar{y}$の不偏標準偏差

$$u_{\bar{y}} = \sqrt{\sum_i \left(\frac{\partial \bar{y}}{\partial y_i}\right)^2 u_y{}^2} = \frac{u_y}{\sqrt{N}}$$

- $\bar{y}$の信頼区間

$$\delta\bar{y} = t_{N-2,\alpha} u_y / \sqrt{N}$$

回帰直線の信頼区間

任意のx_0におけるyの値(y_0)と、その信頼区間を推定する

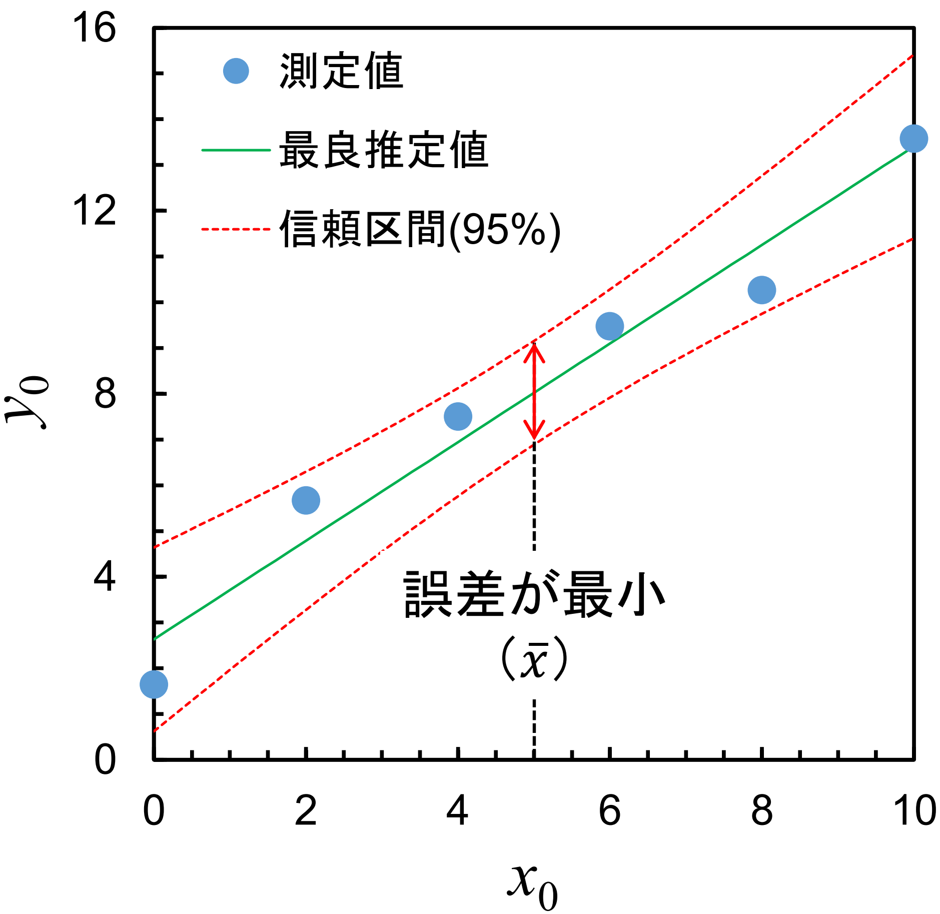
➢y_0 の最良推定値

$$y_0 = ax_0 + b = a(x_0 - \bar{x}) + \bar{y}$$

➢y_0の誤差(aと$\bar{y}$の信頼度は揃える)

$$\delta y_0 = \sqrt{\left(\frac{\partial y_0}{\partial a}\right)^2 (\delta a)^2 + \left(\frac{\partial y_0}{\partial \bar{y}}\right)^2 (\delta \bar{y})^2}$$

$$= \delta y \sqrt{\frac{N}{\Delta}(x_0 - \bar{x})^2 + \frac{1}{N}}$$



- y_0の信頼区間は

$$ax_0 + b \pm \delta y_0$$

逆推定

信号の測定値y_sから未知試料の濃度x_sを推定（検量線法）

最小二乗法で得た検量線: $y = ax + b = a(x - \bar{x}) + \bar{y}$
（y: 信号測定値、x: 濃度）

- 濃度の最良推定値

$$x_s = \frac{y_s - b}{a} = \frac{y_s - \bar{y}}{a} + \bar{x}$$

- 誤差（δy_sとδyの信頼度は揃える）

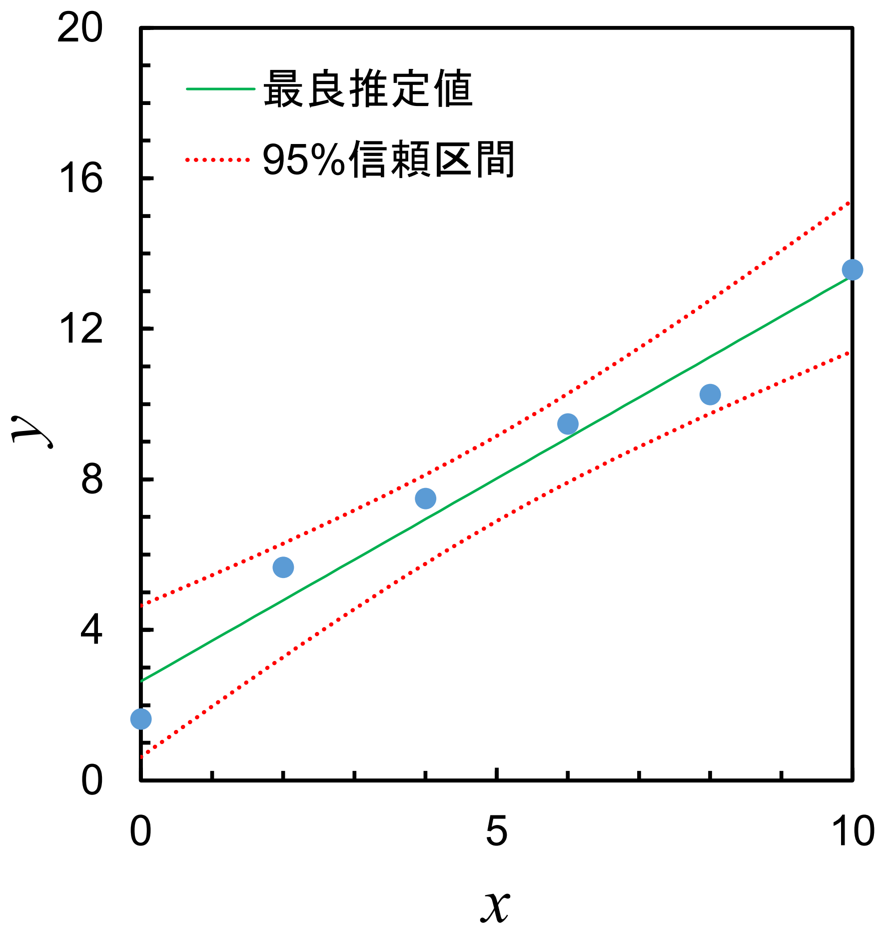
$$\delta x_s = \sqrt{\left(\frac{1}{a}\right)^2 (\delta y_s)^2 + \left[\left(\frac{1}{a}\right)^2 \frac{1}{N} + \left(\frac{y_s - \bar{y}}{a^2}\right)^2 \frac{N}{\Delta}\right] (\delta y)^2}$$

検量線と試料の測定誤差が等しいとみなせる場合

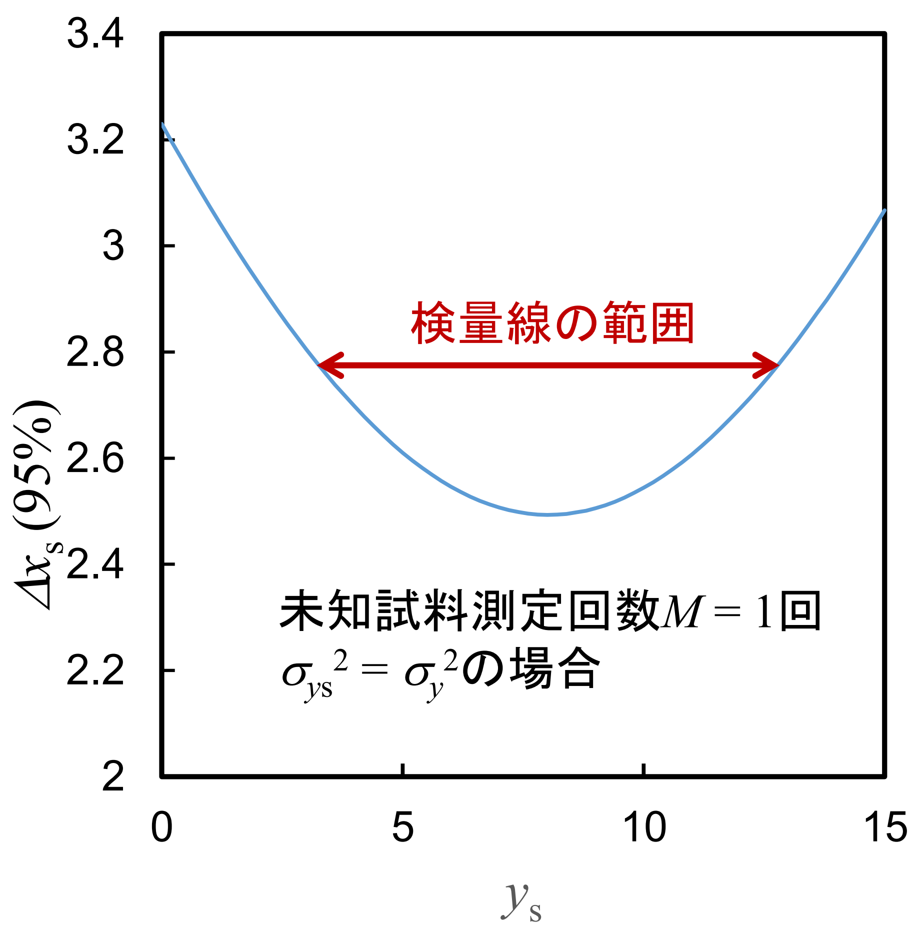
$$\delta x_s = \frac{\delta y}{a} \sqrt{\frac{1}{M} + \frac{1}{N} + \frac{(y_s - \bar{y})^2}{a^2} \frac{N}{\Delta}}$$ （Mは未知試料の測定回数）

逆推定の誤差

➢ 検量線



➢ 誤差

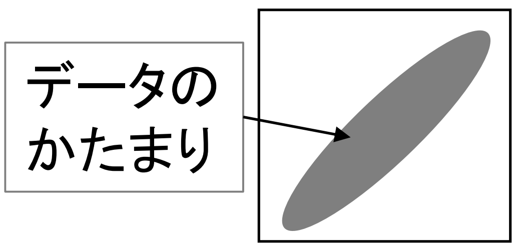


信号の大きさにより誤差が異なる

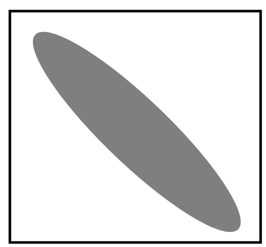
相関係数

- 相関係数r　xとyの間の（線形）相関の指標　$(-1 \leq r \leq 1)$

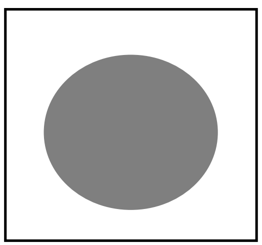
$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}} = \frac{s_{xy}}{s_x s_y}$$



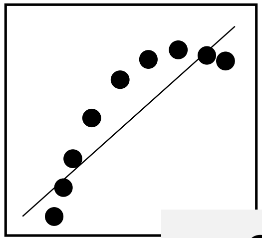
$r>0$
正の相関



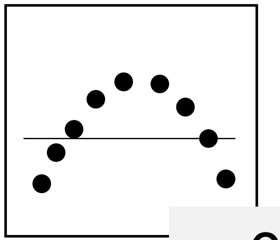
$r<0$
負の相関



$r \sim 0$
相関なし



$r>0$

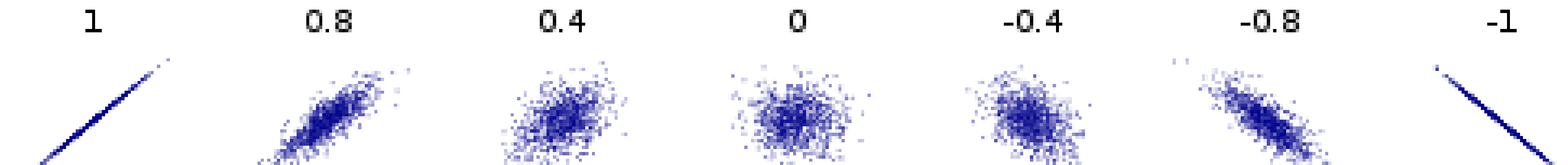


$r \sim 0$

系統誤差はrの値には見えてこないので注意

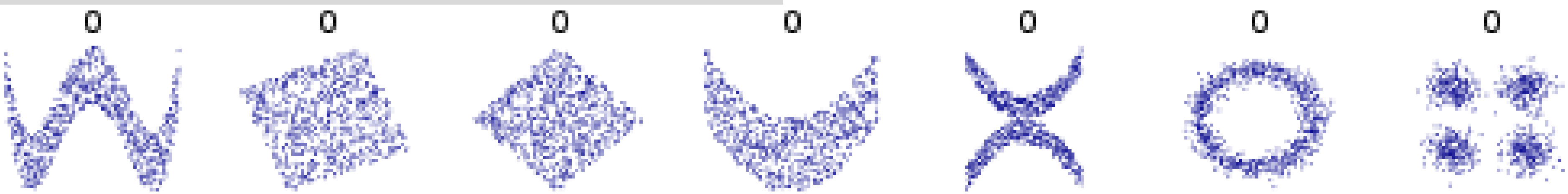
相関係数

データがばらつくほど0に近い



傾きによらない

非線形相関の評価には使えない



(Wikipedia)

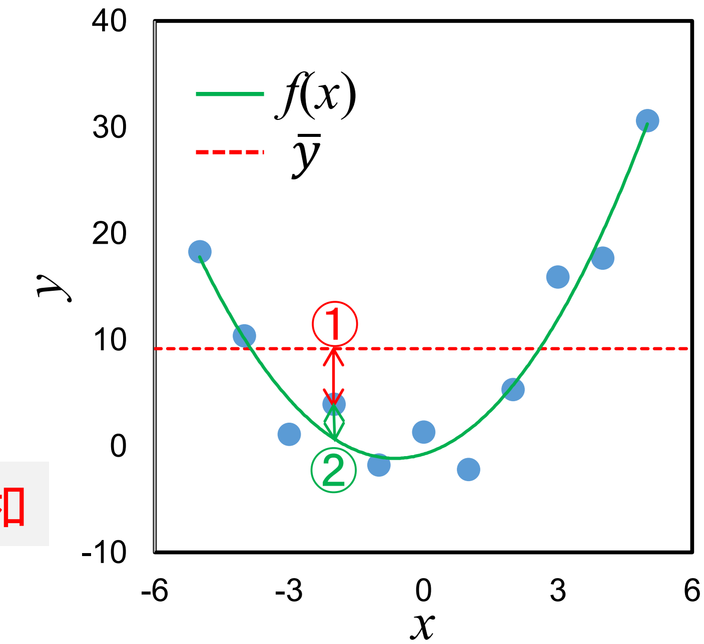
決定係数

➤決定係数R^2

②残差二乗和(RSS)

$$R^2 = 1 - \frac{\sum_i \{y_i - f(x_i)\}^2}{\sum_i \{y_i - \bar{y}\}^2}$$

①y_iの平均まわりの二乗和



R^2が1に近いほど、データとモデル関数が良く合っている

$y=ax+b$による最小二乗法のときは、相関係数rの二乗に等しい　$R^2 = r^2$

モデルの尤もらしさについては評価できない(パラメタが増えればR^2は1に近づく)

相関係数rと同様、系統誤差はR^2の値に見えてこない。
図で誤差のばらつきを確認するべき。

モデルのもっともらしさ(1)

- フィッティングの合う・合わないのひとつの指標は、残差二乗和(RSS)

- だが、パラメタが多いとRSSが下がるのは当たり前

  e.g. *n*個のデータを*n*-1次式でフィットするとRSS=0

  RSSはモデルの尤度(もっともらしさ)の基準にならない

- 赤池の情報量基準AICc　(Akaike's Information Criterion, corrected)

  モデル尤度の基準のひとつ

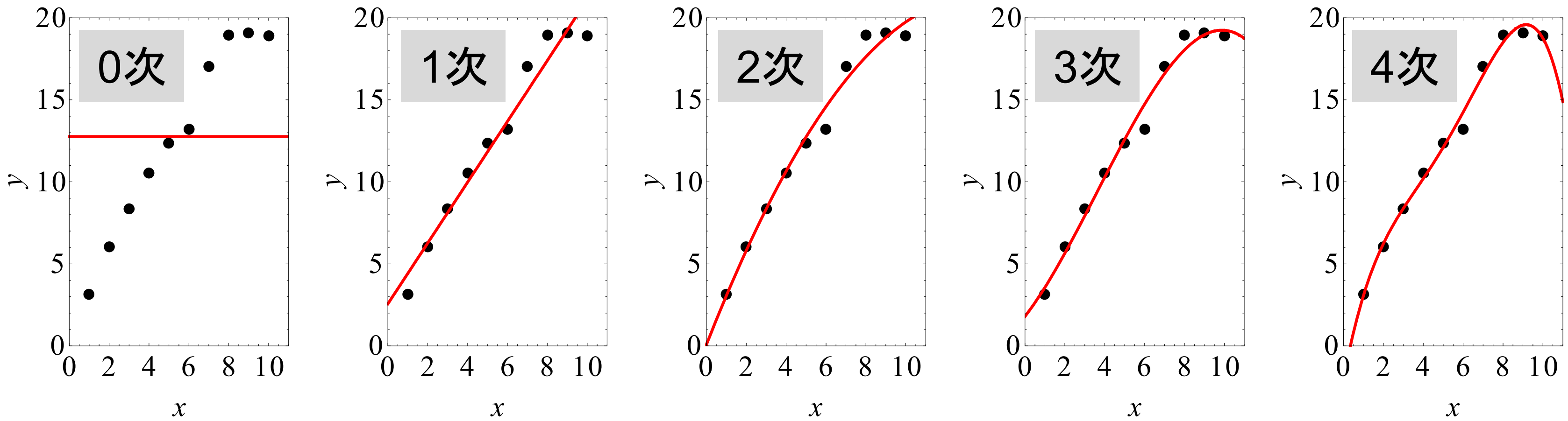
$$\mathrm{AICc} = n \ln \mathrm{RSS} + 2k + \frac{2k(k+1)}{n-k-1}$$

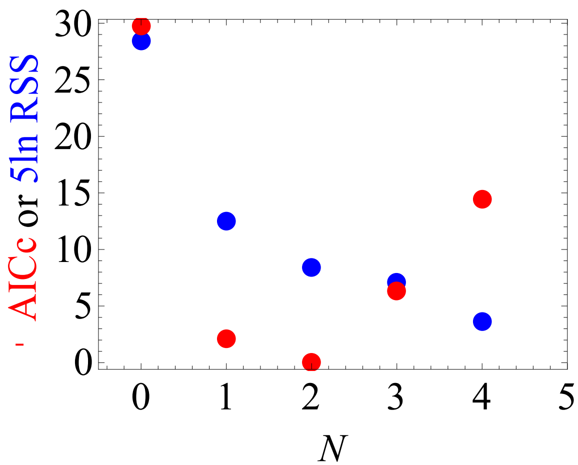
(線形最小二乗法・重みなしのとき, *n*:データ数、*k*:パラメタ数)

AICcが低いほどもっともらしいモデル

モデルのもっともらしさ(2)

- 同一データを次数の異なる多項式でフィッティング





- RSSは次数を上げるほど小さい
  パラメタが増えるので当たり前

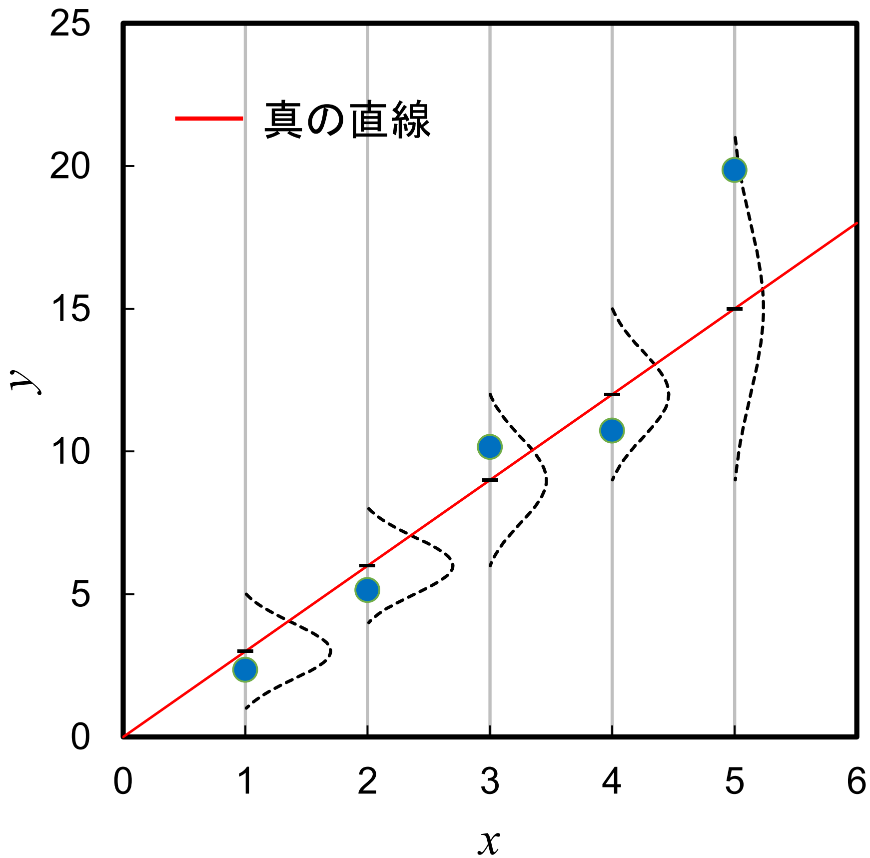
- AICcは二次式(N=2)で最小
  このデータでは二次式モデルが
  最も尤もらしい(統計学的には)

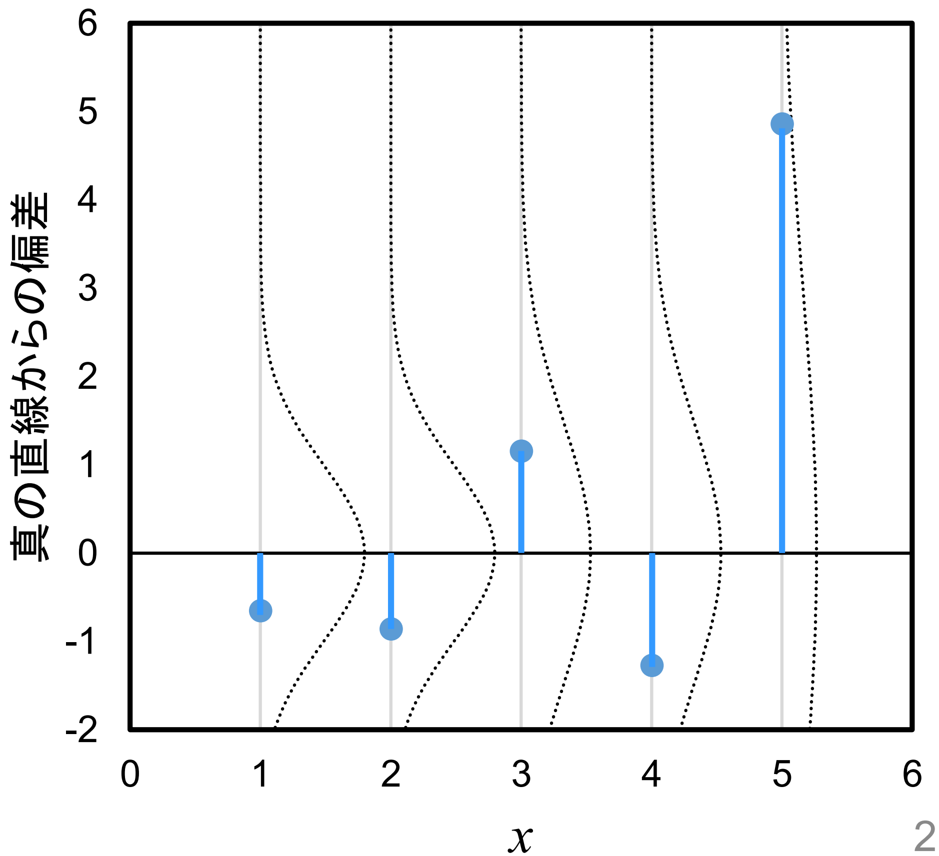
データ間で母分散が異なる場合

重みなし最小二乗法の前提：測定値の母分散は等しい（$\sigma_{yi}{}^2 = \sigma_y{}^2$）

$$\chi^2 = \sum_i \frac{\varepsilon_i{}^2}{\sigma_{yi}{}^2} = \frac{1}{\sigma_y{}^2} \underbrace{\sum_i \varepsilon_i{}^2}_{\text{RSS}}$$

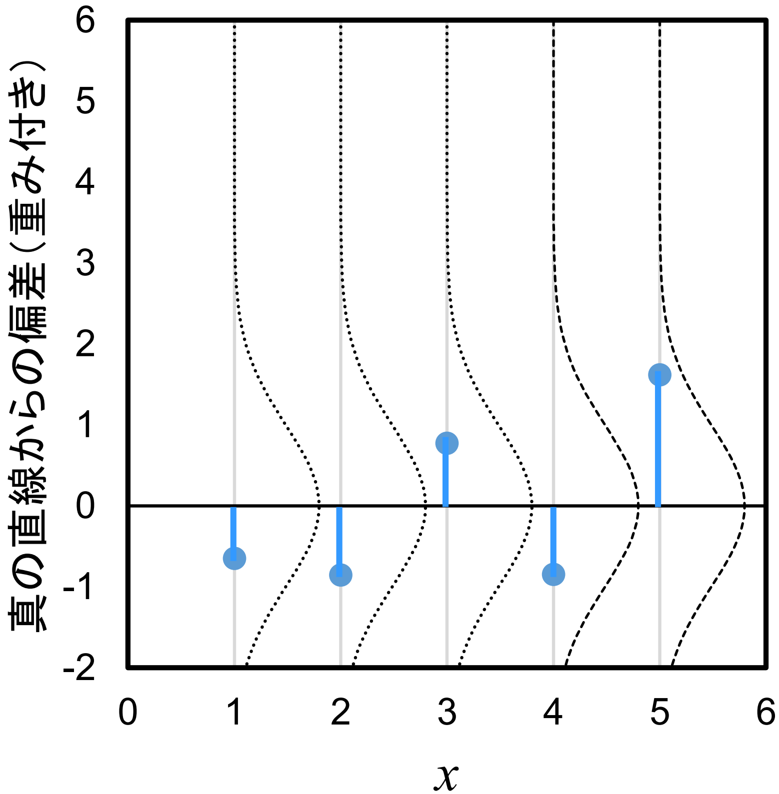
母分散が異なる　⇒　RSSへの寄与が異なる

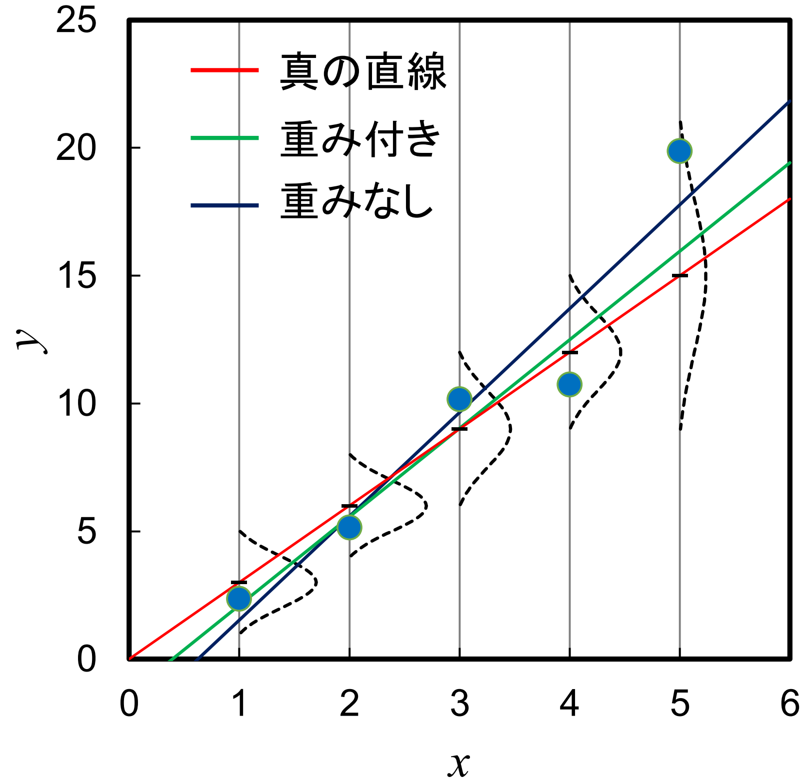




分散の規格化＝重み付け

$$\frac{1}{\sigma_{yi}^{\ 2}} = \frac{w_i}{\sigma_0^{\ 2}} \quad \Rightarrow \quad \chi^2 = \sum_i \frac{\varepsilon_i^{\ 2}}{\sigma_{yi}^{\ 2}} = \frac{1}{\sigma_0^{\ 2}} \underbrace{\sum_i w_i \varepsilon_i^{\ 2}}_{\text{重み付きRSS}}$$





$w_{\mathrm{i}} = 1/\sigma_{yi}^{\ 2}$とすることが多い

重み付けが必要な例

➢ 測定回数が異なる平均値

y_iはn_i回の測定の平均値

$\Rightarrow \sigma_{yi}^2 = \sigma^2 / n_i$　　重み：$w_i = n_i$

$$\chi^2 = \sum_i \frac{\varepsilon_i^2}{\sigma_{yi}^2} = \frac{1}{\sigma^2} \sum_i n_i \varepsilon_i^2$$

➢ 測定値がポアソン分布に従う

度数計測（蛍光計測, MSなど）

$\Rightarrow \sigma_{yi}^2 \propto y_i$　　重み：$w_i = y_i$

$$\chi^2 = \sum_i \frac{\varepsilon_i^2}{\sigma_{yi}^2} \propto \sum_i \frac{\varepsilon_i^2}{y_i}$$

➢ 測定機器が異なる

あるデータはノギスで測定

あるデータはマイクロメータで測定

➢ 測定レンジ（感度）が異なる

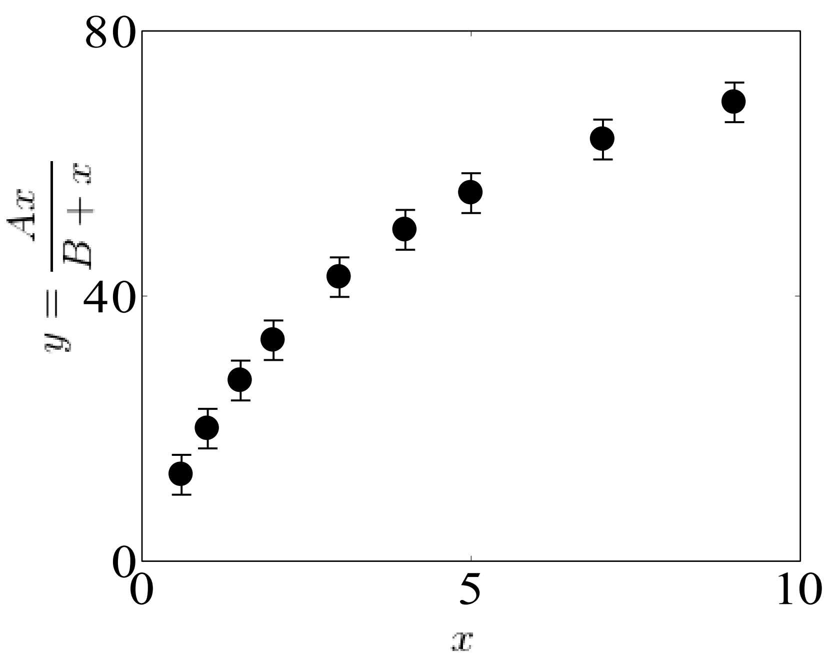
電流の大きさによって
測定レンジを変更

重み付けが必要な例

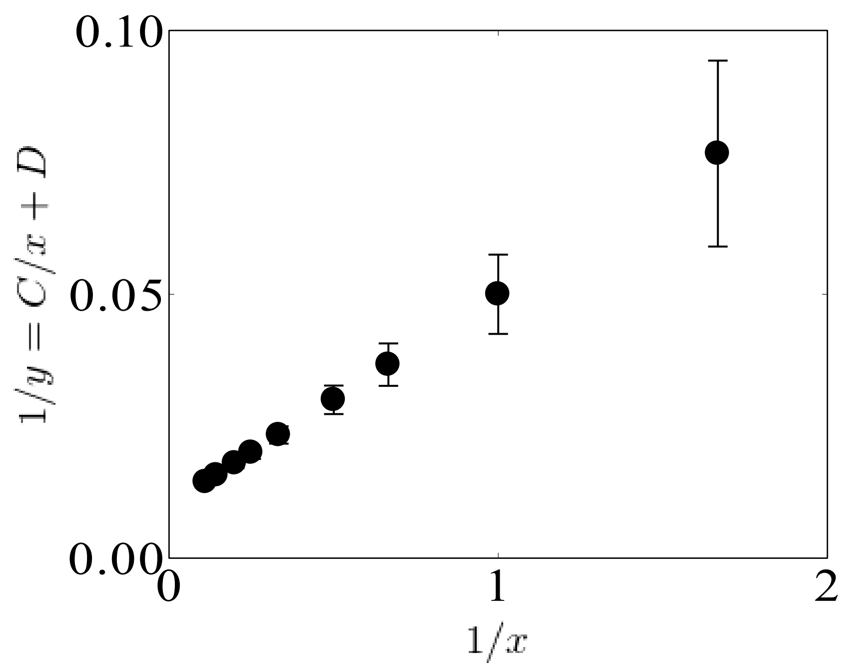
➢データ処理による分散の変化

例：測定値の逆数をプロットする場合

(ミカエリス・メンテン式→ラインウィーバー・バークのプロット)



yのばらつきが均一だとしても



$1/y$のばらつきは不均一

誤差伝播　$\sigma_{1/y}{}^2 = \left(\frac{\partial(1/y)}{\partial y}\right)^2 \sigma_y{}^2 = \frac{1}{y^4}\sigma_y{}^2$　　重み：$w_i = y_i^4$

重み付き線形最小二乗法

- $\Sigma w \varepsilon^2$が最小になるようなパラメタaとbを求めれば良い

$$\frac{1}{\sigma_{yi}^{\ 2}} = \frac{w_i}{\sigma_0^{\ 2}}$$ で重みを定義 (σ_0^2: 規格化後の分散)

$$a = \frac{\Sigma w \Sigma wxy - \Sigma wx \Sigma wy}{\Delta_\omega}$$

$$b = \frac{\Sigma wx^2 \Sigma wy - \Sigma wx \Sigma wxy}{\Delta_\omega}$$

$$\Delta_\omega = \Sigma w \Sigma wx^2 - (\Sigma wx)^2$$

線形最小二乗法の誤差(重み付き)

- 重み付き測定値yの不偏標準偏差u_0

$$u_0 = \sqrt{\frac{\sum_i w_i \varepsilon_i^{\;2}}{N-2}}$$

⇒規格化後の母標準偏差σ_0の最良推定値

- 傾きと切片の不偏標準偏差

δyの誤差伝播から導出

$$u_a = \sqrt{\sum_i \left(\frac{\partial a}{\partial y_i}\right)^2 (u_y)^2} = u_0 \sqrt{\frac{\sum w}{\Delta_w}}$$

$$u_b = \sqrt{\sum_i \left(\frac{\partial b}{\partial y_i}\right)^2 (u_y)^2} = u_0 \sqrt{\frac{\sum wx^2}{\Delta_w}}$$

- 傾きと切片の信頼区間

$$(a-\mu_a)/u_a, (b-\mu_b)/u_b$$

⇒自由度N–2のt分布に従う

$$\delta a = t_{N-2,\alpha} u_a$$

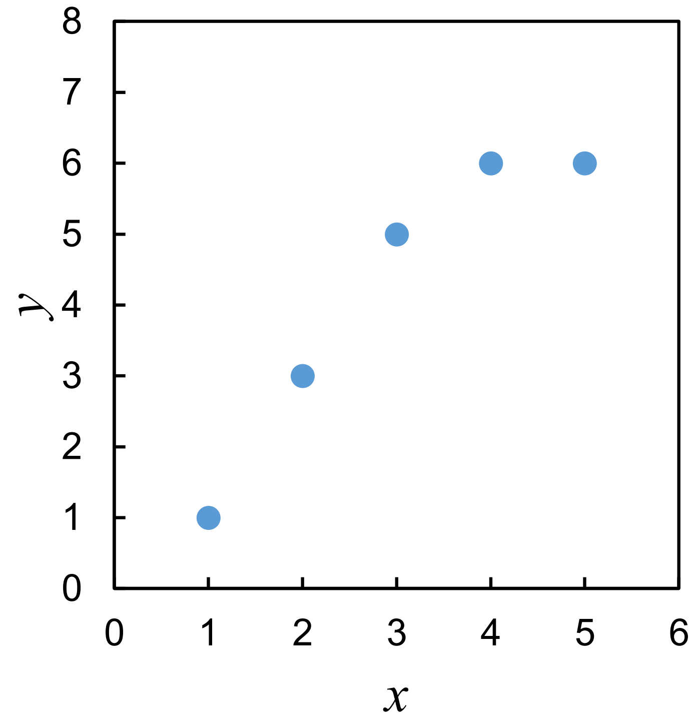
$$\delta b = t_{N-2,\alpha} u_b$$

例題　重み付き最小二乗法

下図のデータにモデル関数$y=ax+b$をあてはめて、パラメタa, bの最良推定値とその95%信頼区間を求めよ。

これらのデータは繰り返し測定の平均値である。一回測定の精度は各データで同じであるが、測定回数n_iが異なる。

x_i	y_i	n_i
1	1	10
2	3	10
3	5	10
4	6	10
5	6	2



例題　重み付き最小二乗法 解説

1.重みを決める。$w_i = n_i$とする。（σ_{yi}が既知なら$1/\sigma_{yi}^2$）

2.測定データからΣw, Σwx, Σwy, Σwx^2, Σwxyを計算

	x_i	y_i	$w_i = n_i$	$w_i x_i$	$w_i y_i$	$w_i x_i^2$	$w_i x_i y_i$
	1	1	10	10	10	10	10
	2	3	10	20	30	40	60
	3	5	10	30	50	90	150
	4	6	10	40	60	160	240
	5	6	2	10	12	50	60
Σ			42	110	162	350	520

3.　正規方程式の解(重み付き)に代入してa, bを計算

$$\Delta_w = 42 \times 350 - (110)^2 = 2600$$

$$a = \frac{42 \times 520 - 110 \times 162}{2600} = \underline{1.55}$$

$$b = \frac{350 \times 162 - 110 \times 520}{2600} = \underline{-0.192}$$

例題　重み付き最小二乗法 解説

4. a, bから重み付きRSS($\Sigma w\varepsilon^2$)を計算

	x_i	y_i	$w_i = n_i$	ax_i+b	$w_i\varepsilon_i^2$
	1	1	10	1.35	1.25
	2	3	10	2.90	0.10
	3	5	10	4.45	3.07
	4	6	10	5.99	0.00
	5	6	2	7.54	4.73
Σ			42		9.15

5. 規格化後の測定値の不偏標準偏差u_0を計算

$$u_0 = \sqrt{\frac{9.15}{5-2}} = 1.75$$

6. u_0からu_a, u_bを計算

$$u_a = 1.75\sqrt{\frac{42}{2600}} = \underline{0.222}$$

$$u_b = 1.75\sqrt{\frac{350}{2600}} = \underline{0.641}$$

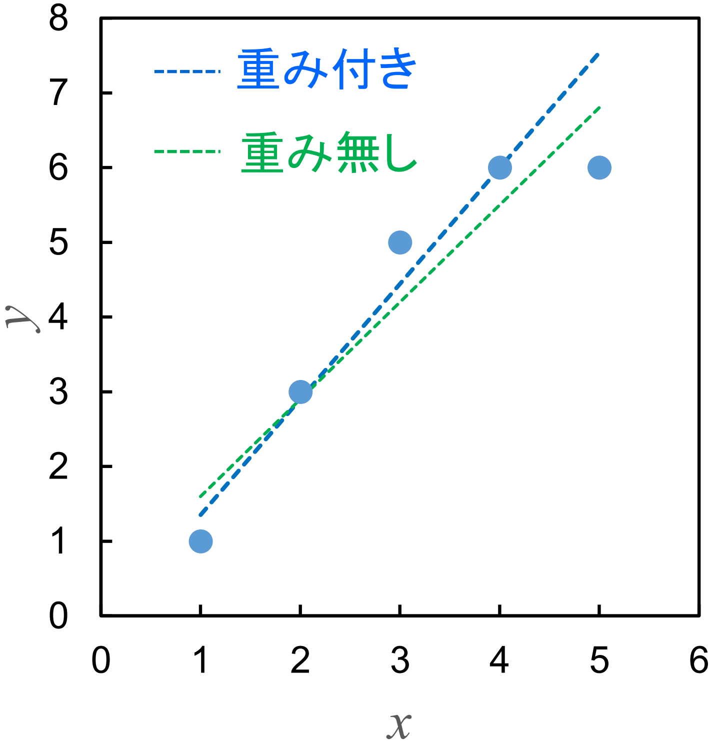
7. それぞれの信頼区間を計算

信頼度α = 95%とすると,

$$\delta a = t_{3,95\%}u_a = \underline{0.62}$$

$$\delta b = t_{3,95\%}u_b = \underline{1.78}$$

$$y = (1.6 \pm 0.6)x + (0 \pm 2)$$



重み無し: $y = (1.3 \pm 0.8)x + (0 \pm 2)$

最小二乗法に関する参考書

- 「実験データを正しく扱うために」 前田・山本・加納
  われらが「緑本」。重み付き線形最小二乗法に詳しい

- 「誤差解析入門」 Taylor
  オススメ。最低限がすべて書かれている
  8章を勉強すれば初級は卒業

- 「最小二乗法の理論と応用」 田島・小牧
  最小二乗法を網羅した分厚い本。三章までの基礎も勉強になる。

- 「最小二乗法による実験データ解析」 中川・小柳
  上級。非線形最小二乗法やAIC、ロバスト推定まで

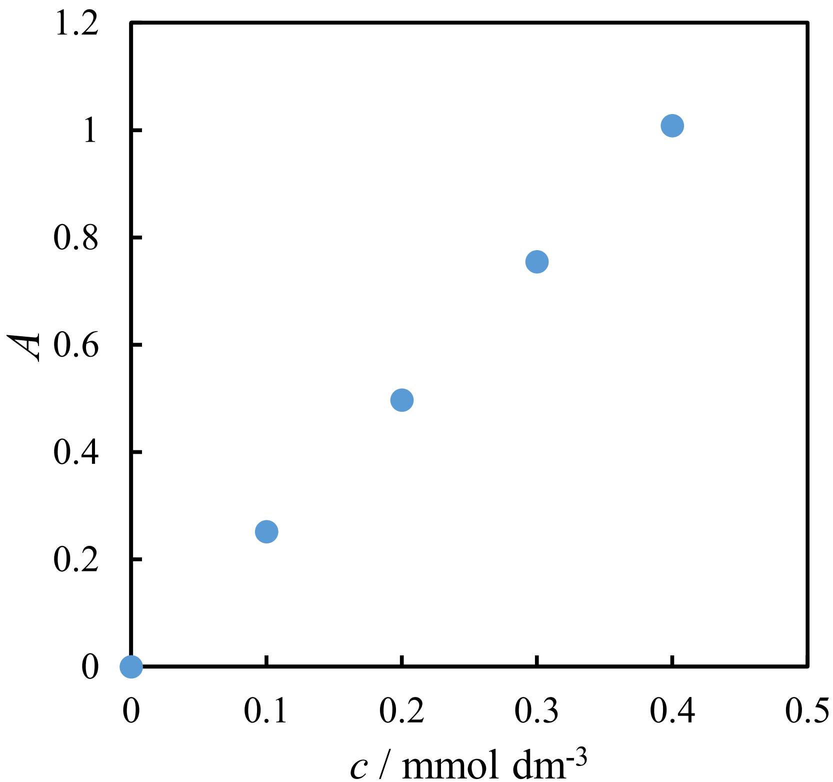
- 「Model Selection and Multimodel Inference」 Burnham・Anderson
  AICにとても詳しい。AICcも。

- 「分析・測定データの統計処理ー分析化学データの扱い方」 田中
  回帰直線の誤差や逆推定について平易に解説

例題　吸光光度法の検量線

化合物Sの520 nmにおけるモル吸光係数を調べるために，様々な濃度(c) の溶液を調製して，光路長 1 cmのセルで吸光度(A)を測定した．以下の測定結果に基づき、モル吸光係数を決定せよ．

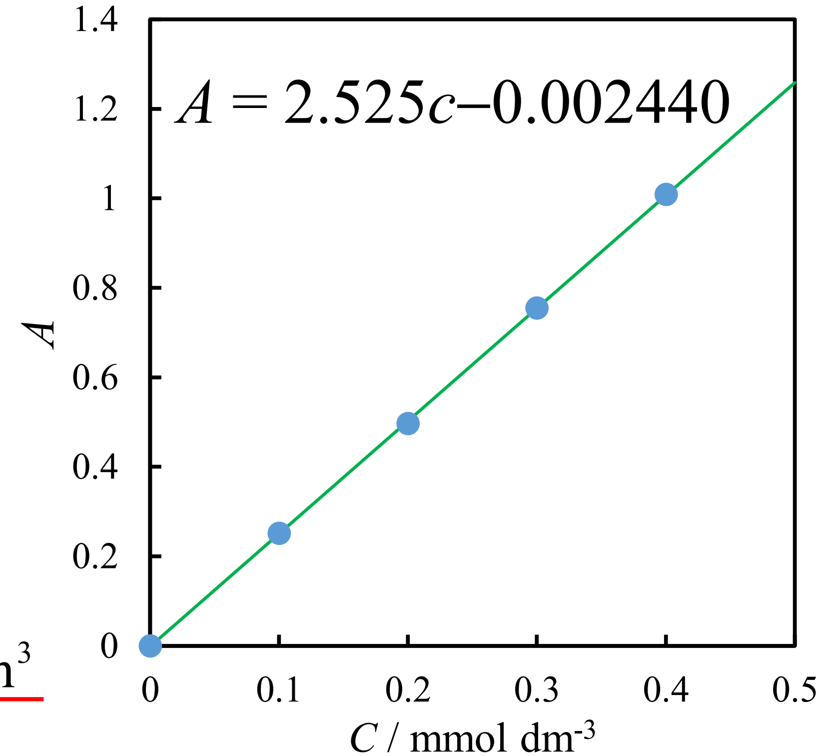
c / mmol dm^{-3}	A / -
0.000	0.0000
0.100	0.2517
0.200	0.4970
0.300	0.7553
0.400	1.0086



例題　吸光光度法の検量線　解説

1. 測定データからΣx, Σy, Σx^2, Σxyを計算

	c / mM x_i	A / - y_i	x_i^2	$x_i y_i$
	0.000	0.0000	0.0000	0.0000
	0.100	0.2517	0.0100	0.0252
	0.200	0.4970	0.0400	0.0994
	0.300	0.7553	0.0900	0.227
	0.400	1.0086	0.160	0.403
Σ	1.000	2.5126	0.3000	0.755



2.　正規方程式の解に代入してa, bを計算

$$\varDelta = 5\times 0.3000 - (1.000)^2 = 0.500\ \mathrm{mmol^2\ dm^{-6}}$$

$$a = \frac{5\times 0.755 - 1.000\times 2.5126}{0.500} = \underline{2.525\ \mathrm{mmol^{-1}dm^3}}$$

$$b = \frac{0.300\times 2.5126 - 1.000\times 0.755}{0.500} = \underline{-0.002440}$$

例題　吸光光度法の検量線　解説

3. a, bからRSS（$\Sigma\varepsilon^2$）を計算

	c /mM x_i	A y_i	ax_i+b	ε_i^2 / 10^{-6}
	0.000	0.0000	-0.002	2.690
	0.100	0.2517	0.2501	1.588
	0.200	0.4970	0.5022	30.05
	0.300	0.7553	0.7543	0.490
	0.400	1.0086	1.0064	3.686
Σ	1.000	2.5126		38.92

4. 測定値の標準偏差u_yを計算

$$u_y = \sqrt{\frac{38.92\times10^{-6}}{5-2}} = \underline{3.602\times10^{-3}}$$

5. u_yからu_a, u_bを計算

$$u_a = 3.602\times10^{-3}\sqrt{\frac{5}{0.500}} = \underline{1.13\times10^{-2}\ \mathrm{mM}^{-1}}$$

$$u_b = 3.602\times10^{-3}\sqrt{\frac{0.300}{0.500}} = \underline{2.79\times10^{-3}}$$

6. それぞれの信頼区間を計算

信頼度α = 95%とすると,

$$\delta a = t_{3,95\%}u_a = \underline{3.6\times10^{-2}\ \mathrm{mM}^{-1}}$$

$$\delta b = t_{3,95\%}u_b = \underline{8.9\times10^{-3}}$$

$$A = (2.52 \pm 0.04)c + (-0.002 \pm 0.009)$$

モル吸光係数は

$$(2.52 \pm 0.04) \times 10^4\ \mathrm{M}^{-1}\ \mathrm{cm}^{-1}$$

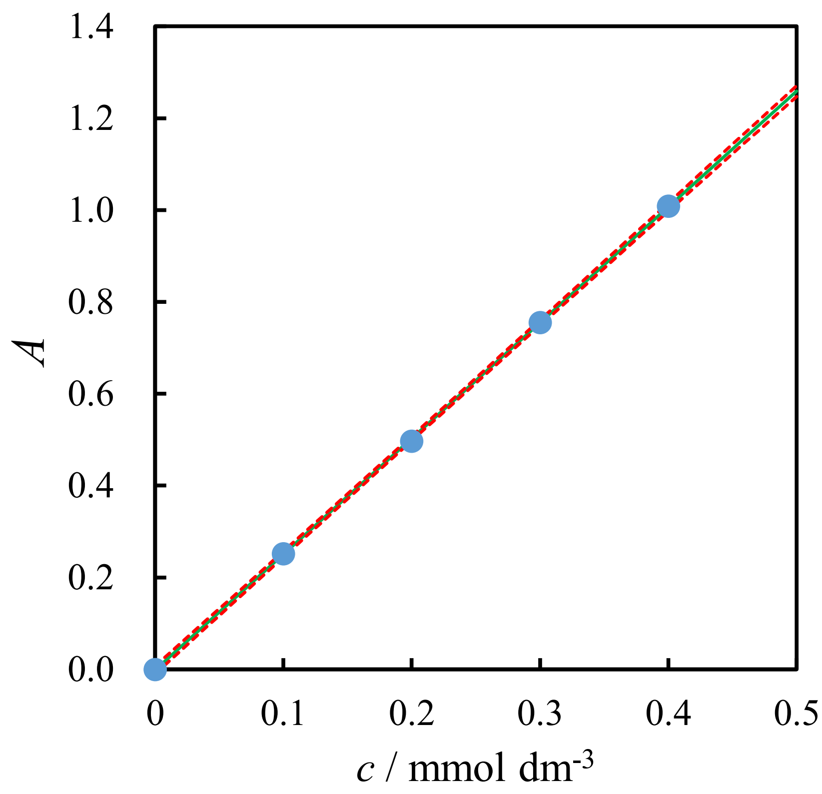
例題　逆推定

化合物S溶液（濃度未知）の520nmにおける吸光度を1回測定したところ、0.600であった（1cmセル中）。

問2-1で得た濃度と吸光度の関係を検量線として、この溶液中の化合物*S*の濃度を求めよ。

なお、この溶液の組成は、検量線用の標準溶液と化合物*S*の濃度以外同じである。

$A = (2.52 \pm 0.04)c + (-0.002 \pm 0.009)$



例題　逆推定　解説

化合物Sの濃度

$$x_s = \frac{y_s - b}{a} = \frac{0.6000 - (-0.002440)}{2.525} = 0.2386\,\mathrm{mM}$$

誤差

- 未知試料と標準溶液の測定誤差は一定
- 吸光度の測定は1回

$M = 1,\ N = 5,\ \bar{y} = 0.502,\ \delta y = 1.15\times10^{-2}$ (信頼度95%)

$$\delta x_s = \frac{\delta y}{a}\sqrt{\frac{1}{M} + \frac{1}{N} + \frac{(y_s - \bar{y})^2}{a^2}\frac{N}{\Delta}}$$

$$= \frac{1.15\times10^{-2}}{2.525}\sqrt{\frac{1}{1} + \frac{1}{5} + \frac{(0.600 - 0.502)^2}{2.525^2}\frac{5}{0.500}} = 5.02\times10^{-3}\ \mathrm{mM}$$

答　$(2.39 \pm 0.05) \times 10^{-4}\ \mathrm{mol\ dm^{-3}}$

27th May, 2017

不確かさの概念と無機分析における見積もり例

0011 0010 1010 1101 0001 0100 1011

明星大学大学院理工学研究科　上本道久

1. 不確かさと誤差の違い
2. 不確かさの考え方
3. 無機分析における見積もり例

– ICP発光分析法による定量分析の場合

２．不確かさとは

不確かさ(uncertainty)；測定結果に対する信頼度に関した、統一的な計測の信頼性表現として、ISOを含む複数の国際機関が連携して提案した用語

不確かさの定義

- 「測定量の真の値が存在する範囲を示す推定値」

  JIS Z 8103「計測用語」(旧規格、1990)

VIM（International vocabulary of basic and general terms in metrology–国際計量基本用語集）第１版　に準拠

- 「測定の結果に付記される、合理的に測定量に結び付けられ得る値のばらつきを特徴づけるパラメーター」

  JIS Z 8103「計測用語」(2000改訂)

VIM第２版およびGUM（Guide to the expression of uncertainty in measurement–計測における不確かさの表現ガイド）第１版(1993)に準拠

信頼性の定義

- 「精度又は正確さの期待できる程度」

  JIS K　0211「分析化学用語(基礎部門）」

  (旧規格、1987)

- 「機器、方法又はそれらの要素が、規定の条件の範囲内において規定の機能と性能を保持する性質又は度合い」

  JIS K　0211「分析化学用語(基礎部門）」

  (2013改訂)

２．不確かさの考え方

★不確かさの内容を調べて評価すると、一例として以下のようなことがわかる。

１）分析技術者の熟練度

２）装置、器具の適否

３）分析法の適否

４）試薬、純水の純度などの適否

５）分析室を取り巻く環境の適否

★各誤差要因の標準不確かさを求めて、次式により合成して、合成標準不確かさを算出する。

$$u(y(x_1,x_2,....))=\sqrt{\sum_{i=1,n} c_i^2\, u(x_i)^2}$$

C_iは感度係数で、x_iの関数yの偏微分として表されるが、実験的にも直接求められる。

$$c_i = \partial y / \partial x_i$$

「不確かさの見積もりにあたっての考え方

★不確かさの評価は型にはまった作業でもなく、純数学的なものでもない。不確かさ要因を識別するのは実験者である。

★かたよりについては、あらかじめ評価できるものは補正する。

★不確かさは小さければ良いという単純なものではない。

統計的考察の前に分析化学的考察を行うべし

Ｚスコアに関わる問題

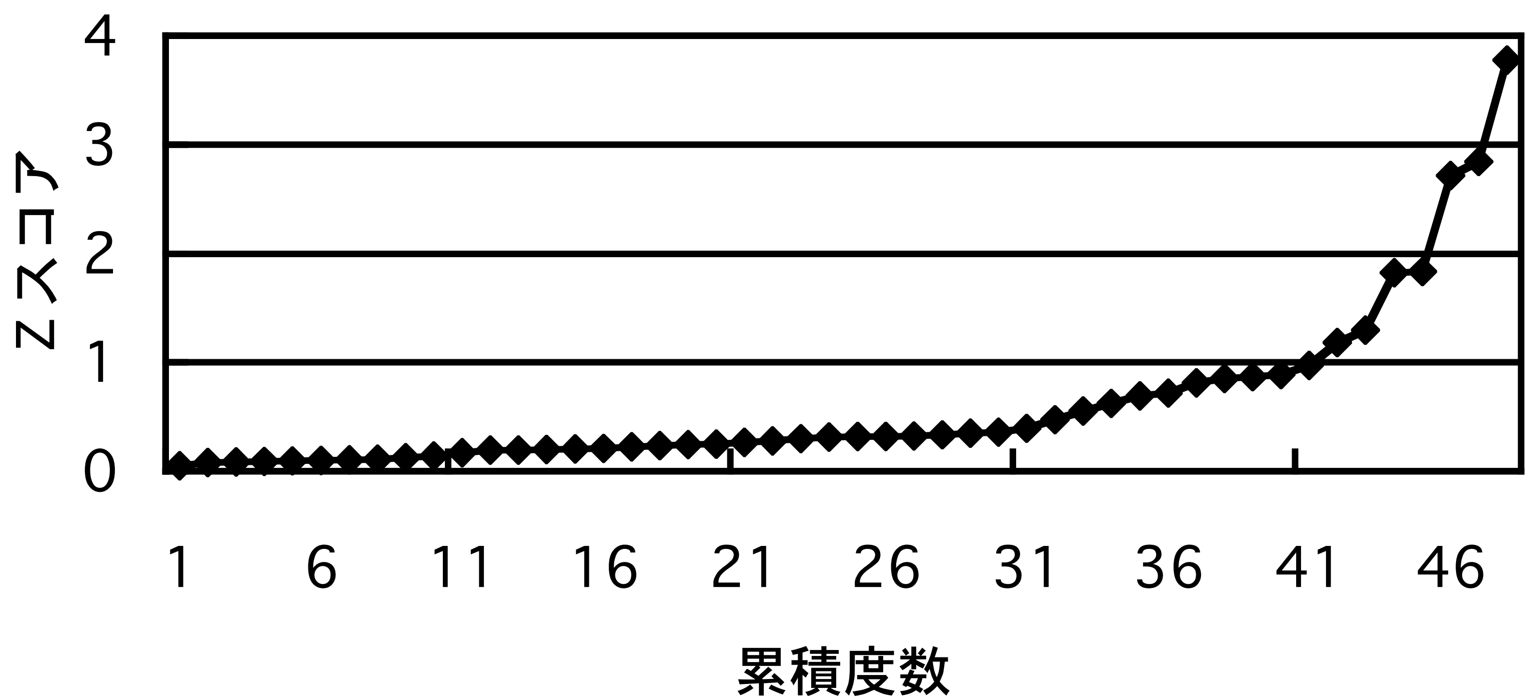
$$Z = \frac{|（測定値）－（平均値）|}{（母集団の標準偏差）}$$

★平均値に近いほどＺスコアは小さくなる。

「多頻度の数値に依存する」

★平均値は'真値'を反映しているという前提がある。

★ばらつきが大きければ（分析値の質が悪ければ）Zスコアは小さくなる。

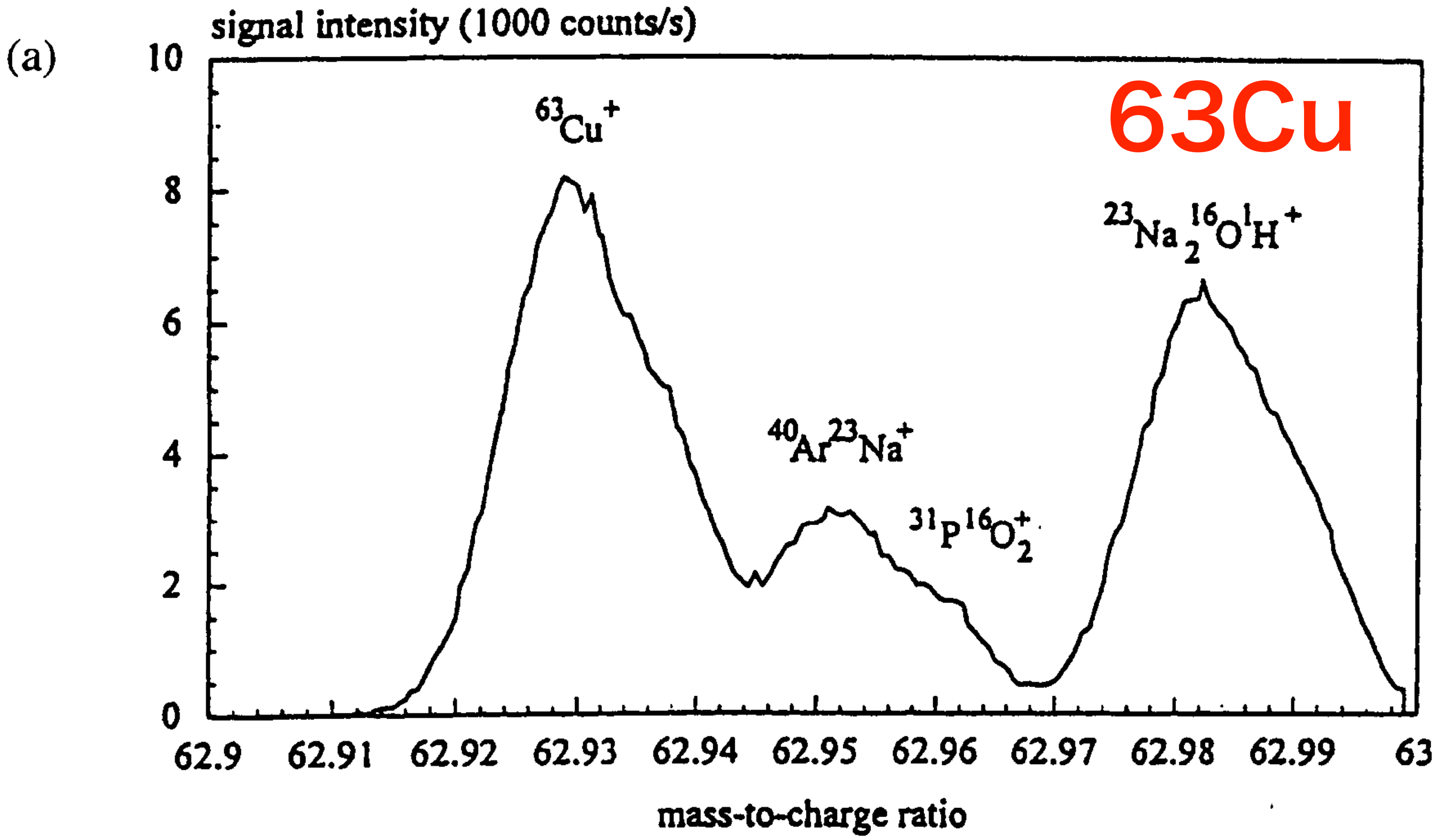


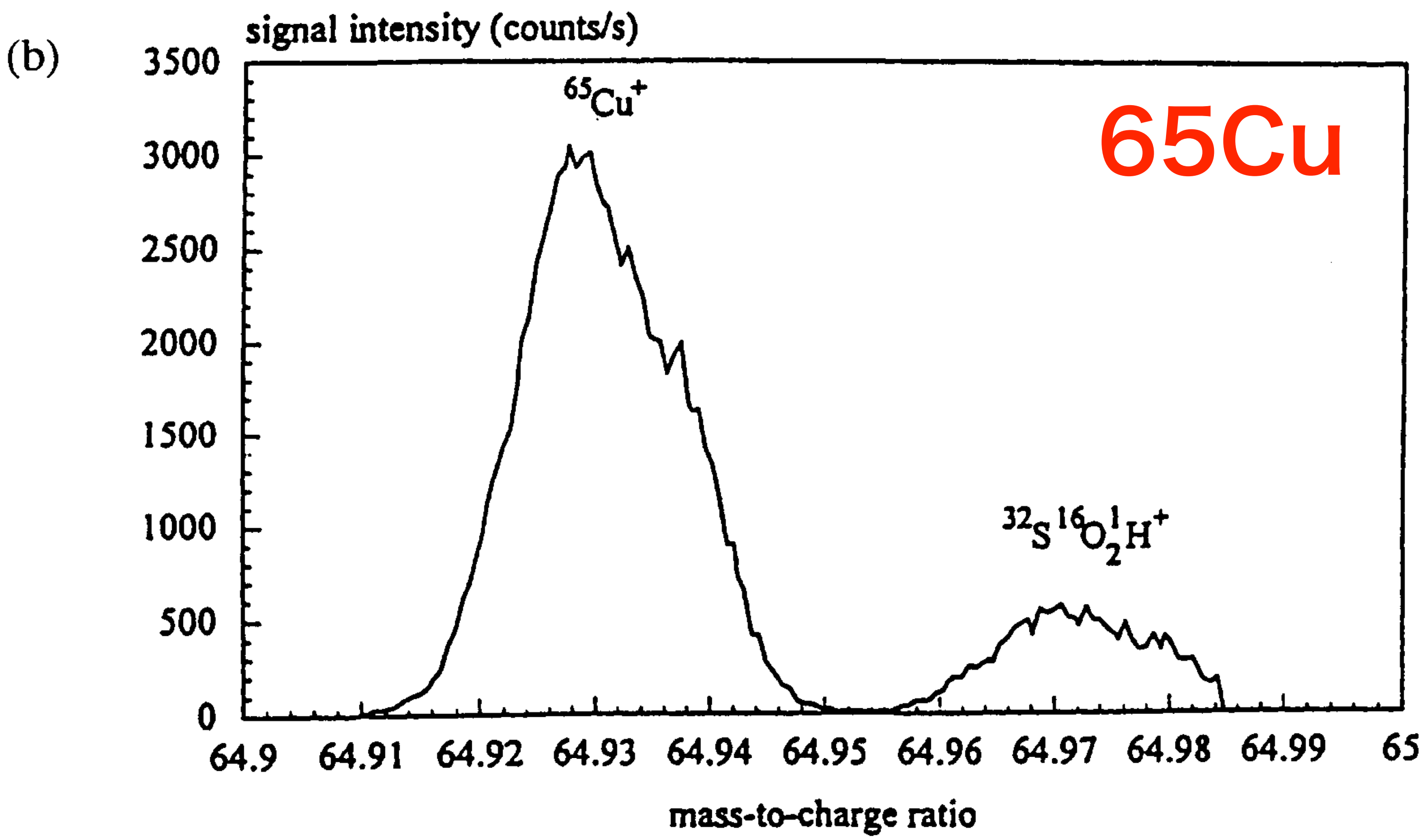
共同分析における
ｚスコアの一例

Alを約100μg cm^{-3}を含む試料中のNaを原子スペクトル分析により定量する　ー共同分析結果の一例

- 標準溶液に試料と同程度の濃度の主成分元素を添加して干渉を抑える操作（マトリックスマッチング、MM）が必要である。
- 測定対象元素だけを含む標準溶液で検量線を引き、定量する（MM無）分析者が意外に多い。

	度数	平均	S.D.	C.V.(%)
＜全体＞				
MM有	18	5.94	0.312	5.3
MM無	37	6.11	0.827	13.5
割合(%)	33			





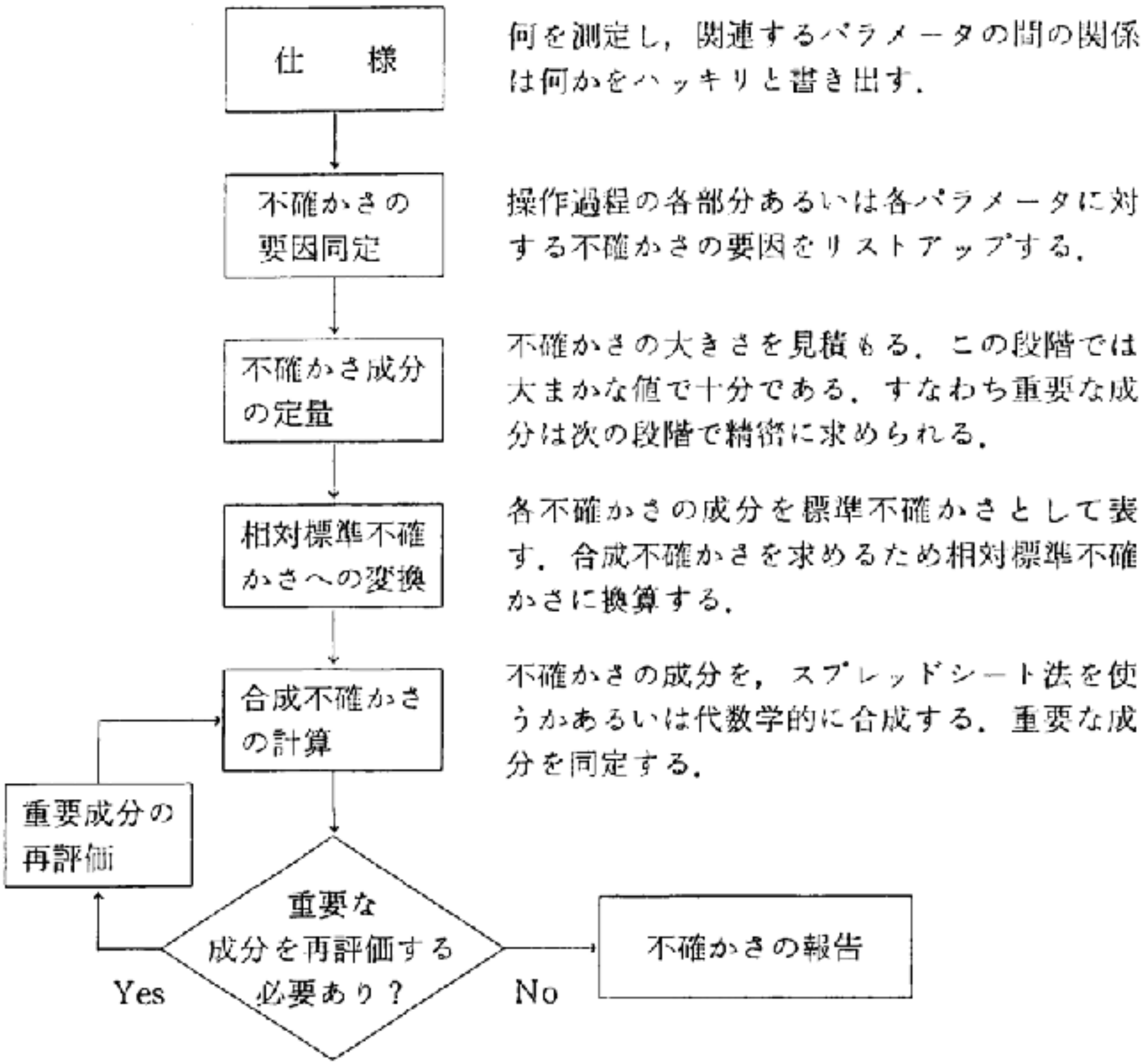
血清を希釈して硝酸酸性水溶液とした試料のHR-ICP-MSスペクトル

cf. S.J.Hill ed., Inductively Coupled Plasma Spectrometry and its Applications, (CRC Press, 1997),p.190

両同位体共、分子イオンのスペクトル干渉を受ける

↓

四重極型ICP-MSでは、希釈だけで極微量のCuを定量することは出来ない



不確かさの見積もりのプロセス

不確かさの見積もりの例

「ICP発光分析により試料中の微量金属元素を定量する」

1）<u>**試料処理**</u>

　秤量、溶解、定容

2）<u>**標準溶液の調製**</u>

　市販標準溶液から段階的に希釈

3）<u>**ICP発光分析装置を用いた測定**</u>

　標準溶液を測定ー検量線の作成

　試料溶液の測定ー試料の測定値の算出

不確かさ：測定の結果に付記される、合理的に測定量に結び付けられ得る値の<u>ばらつきを特徴づける</u>パラメーター

・最も簡単な不確かさの見積もり

独立したサンプリングによるｎ回の分析（n連）の分析値の平均と標準偏差がそれぞれｘ,σのとき、数値に偏りがなければ、分析値の不確かさは以下で与えられる。

分析値： $x \pm \sigma/\sqrt{n}$

①不確かさ要因の洗い出しと同定

★統計的に評価し得るタイプＡと呼ばれる要因（偶然誤差に相当）

★それ以外により評価されるタイプＢと呼ばれる要因（系統誤差に相当）

②それぞれの要因（成分）の標準不確かさの算出

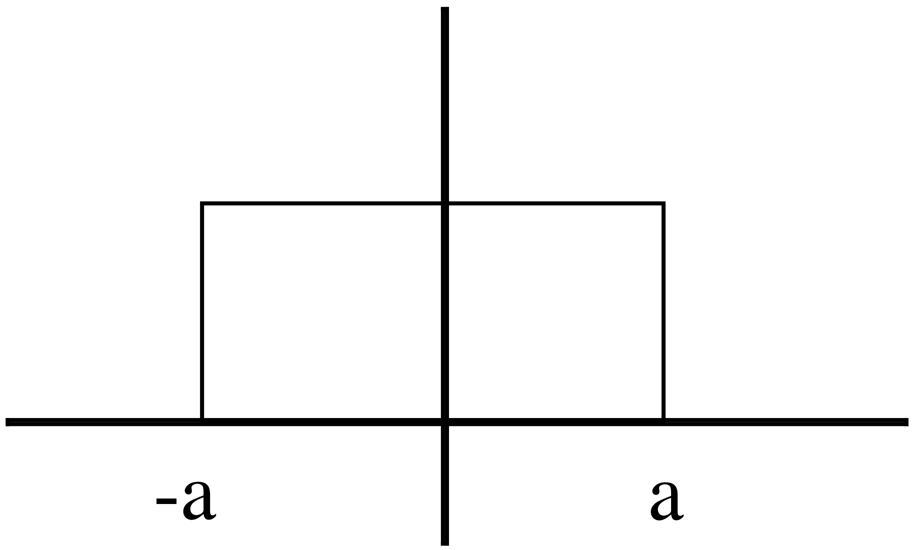
タイプＡ　：単純にｎ回の繰り返し測定の標準偏差として求められる。また、一つの母集団から取られたm個の試料の平均値の標準偏差は、一つの試料をｎ回測定した標準偏差（前者）を√mで割った値として与えられる。

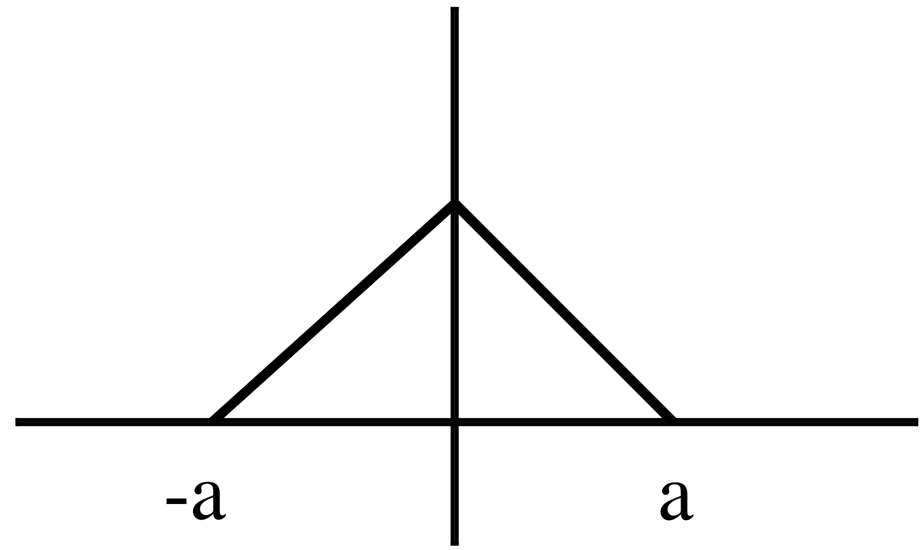
タイプＢ　：信頼性レベルの表示がなく、極端な値があるかもしれない状況下で±ａという範囲が与えられていれば、その標準不確かさはａ/√３とする。信頼性レベルの表示がないが、極端な値がないという理由がある場合は、±ａという範囲が与えられていれば、その標準不確かさはａ/√６とする。

③全ての成分を加味した合成標準不確かさの算出

④合成標準不確かさに包含係数を掛けた拡張不確かさの算出

分布関数





矩形分布

$$u = \frac{a}{\sqrt{3}}$$

三角分布

$$u = \frac{a}{\sqrt{6}}$$

試料処理

・試料の秤量;0.5023 g

(Aタイプ)無視

(Bタイプ)セミミクロ電子天秤の直線性など:

0.0002 g

$u_{C1} = 0.0002/\sqrt{3} = 0.00012$

$u_{C1,r} = 0.00012/0.5023 = 0.00024$

・酸で溶解後、100 mlに調製

全量(メス)フラスコ使用

全量(メス）フラスコ
100 ml

- Aタイプ；0.05
- Bタイプ；$0.1/\sqrt{3}=0.058$
- 合成標準不確かさ；

  $\sqrt{(0.05^2+0.058^2)}=0.077$
- 相対合成標準不確かさ；

  $u_{c2,r}= 0.077/100 = 0.00077$

不確かさ見積もり手順

- ①ばらつきに関して、測容器で繰り返し(10回以上）純水を計り取りその質量を測定する。(タイプA）
  純水の密度より体積に換算し、繰り返し再現性に起因する標準不確かさを見積もる。
- 秤量びんに10mlのWP から水を移して秤量したところ10.0435 g
  →10.0435/ 0.99820 (20℃の水の密度)= 10.062 ml
- ②測容器の公差に起因する標準不確かさを見積もる。
  (タイプB）
- ③前二者を加味した合成標準不確かさを算出する。

全量ピペットにおける採取重量の繰り返し測定結果とその時の室内温度

回数	1ml WP		2ml WP		3ml WP		5ml WP		10ml WP	
	水の重量 g	温度 T/℃	水の重量 g	温度 T/℃	水の重量 g	温度 T/℃	水の重量 g	温度 T/℃	水の重量 g	温度 T/℃
1	0.9973	25.9	2.0062	25.9	2.9938	25.9	4.9753	25.9	9.9645	25.9
2	0.9981	25.9	2.0084	25.9	2.9828	25.9	4.9758	25.9	9.9570	25.9
3	0.9973	25.9	2.0065	25.9	2.9882	25.9	4.9691	25.9	9.9585	25.9
4	0.9984	25.9	2.0110	25.9	2.9993	25.9	4.9755	25.9	9.9616	25.9
5	0.9992	25.9	2.0023	25.9	2.9854	25.9	4.9798	25.9	9.9546	25.9
6	0.9976	25.9	2.0042	25.9	2.9889	25.9	4.9733	25.9	9.9524	25.9
7	0.9971	25.9	2.0005	25.9	2.9862	25.9	4.9845	25.9	9.9643	25.9
8	1.0001	25.9	2.0052	25.9	2.9865	25.9	4.9755	25.9	9.9594	25.9
9	1.0068	25.9	2.0054	25.9	2.9871	25.9	4.9889	25.9	9.9479	25.9
10	0.9959	25.9	2.0026	25.9	2.9895	25.9	4.9951	25.9	9.9668	25.9

測容器の合成不確かさ

$$Uc = \sqrt{\{u_1(\text{再現性})^2 + u_2(\text{公差}/\sqrt{3})^2\}}/v$$

全量ピペットの容積	平均体積 /ml	再現性 (S.D.)	相対標準偏差 RSD (%)	公差/√3	合成標準不確かさ Uc(RSD)
1ml WP	1.0020	0.0031	0.304	0.01	0.0065
2ml WP	2.0117	0.0031	0.153	0.01	0.0033
3ml WP	2.9984	0.0047	0.158	0.01	0.0025
5ml WP	4.9953	0.0080	0.160	0.02	0.0028
10ml WP	9.9908	0.0059	0.060	0.02	0.0026

全量フラスコの容積と材質	平均体積 /ml	再現性 (S.D.)	相対標準偏差 RSD (%)	公差/√3	合成標準不確かさ Uc(RSD)
25ml MF	24.9874	0.0204	0.081	0.06	0.0016
50ml MF	50.0284	0.0206	0.041	0.10	0.0012
100ml MF	99.9985	0.0221	0.022	0.12	0.0007
PP製 100ml MF	100.5158	0.0233	0.023	0.12	0.0007
PMP製 100ml MF	100.6856	0.0349	0.035	0.12	0.0008

2) 標準溶液の調製

- ・1000 g cm^{-3}の市販金属標準液（国産品は通常1000± 6 g cm^{-3}と記載, k=2であることが多いが、ここではそのまま使う）

$u_{std}=6/\sqrt{3}=3.5$

- ・ 1000 g cm^{-3}標準液から段階的に希釈して、0, 0.1, 0.2, 0.4 g cm^{-3}の4つの標準溶液を調製

容器での希釈操作における不確かさ

濃度が1000 μgcm^{-3}の標準原液を、検量線作成のために４段階で10,000倍に希釈したと仮定する。

１段階；1000 μgcm^{-3} ; → 100 μgcm^{-3}

(10 cm^{3}のホールピペット／100 cm^{3}のメスフラスコ)

２段階； 100 μgcm^{-3} ; → 5 μgcm^{-3}

(5 cm^{3}のホールピペット／100 cm^{3}のメスフラスコ)

３段階； 5 μgcm^{-3} ; → 0.5 μgcm^{-3}

(10 cm^{3}のホールピペット／100 cm^{3}のメスフラスコ)

４段階；0.5 μgcm^{-3} ; → 0.1 μgcm^{-3}

(20 cm^{3}のホールピペット／100 cm^{3}のメスフラスコ)

測容器の合成不確かさ

器具	再現性(SD)	(公差)／√3	u c (RSD)
5 cm^3 WP	0.0048	0.015/√3=0.0087	$\{\sqrt{(0.0048^2+0.0087^2)}\}/5$ =0.0020
10 cm^3 WP	0.012	0.02/√3=0.012	$\{\sqrt{(0.012^2+0.012^2)}\}/10$ = 0.0017
20 cm^3 WP	0.03	0.03/√3=0.017	$\{\sqrt{(0.03^2+0.017^2)}\}/20$ = 0.0017
100 cm^3 MF	0.05	0.1/√3=0.058	$\{\sqrt{(0.05^2+0.058^2)}\}/100$ = 0.00077

全量（ホール） ピペット
10 ml

u_c (RSD) =0.0017

全量(メス）フラスコ
100 ml

u_c (RSD) =0.00077

10ml 採取して100mlにメスアップ → 10倍希釈

u_c (RSD) $=\sqrt{(0.0017^2+0.00077^2)} = 0.0019$

各希釈操作における合成不確かさ

器具	uc（RSD)
１段階（10倍希釈）	；$\sqrt{}(0.0017^2+0.00077^2) = 0.0019$
２段階（20倍希釈）	；$\sqrt{}(0.0020^2+0.00077^2) = 0.0021$
３段階（10倍希釈）	；$\sqrt{}(0.0017^2+0.00077^2) = 0.0019$
４段階（5倍希釈）	；$\sqrt{}(0.0017^2+0.00077^{22}) = 0.0019$

4段階の希釈により、計10,000倍に希釈したとすれば、

$u_C = \sqrt{}(0.0019^2+0.00021^2+0.0019^2+0.0019^2)$

$= 0.0039$ (RSD)

標準溶液調製における相対合成標準不確かさ

0.1 μgcm^{-3} ; 1000 μgcm^{-3}を4段階で希釈

$$u_{c3,r} = \sqrt{\left(\frac{3.5}{1000}\right)^2 + 0.0039^2} = 0.0052$$

0.1, 0.2, 0.4 mg cm-3のうち、最も希釈率の大きい低濃度溶液が最も不確かさが大きいと考えられる

→0.0052を標準溶液調製における相対合成標準不確かさとする

検量線法による相対分析

- 被測定物質濃度 x、信号強度 y、
- 検量線 y=ax +b（aとbは最小二乗法にて決定）
- 標準液測定点数n、試料溶液の繰返し測定m回
- xおよびyの平均値：x_{ave}、y_{ave}

残差標準偏差の平方s_y^2:

$$s_y^2=\Sigma\{(ax_i+b)-y_i\}^2/(n-2)$$

通常の検量線法で計算した試料濃度x_Aに対応する標準不確かさ$u(x_A)$は

$$u^2(x_A)=s_y^2/a^2\,[1/m + 1/n + (y_A-y_{ave})^2/a^2\,\Sigma(x_A-x_{ave})^2\,]$$

一例として、測定値0.254 gcm^{-3}に対して
$u(x_A)=u_{C4}=0.0029$ gcm^{-3}と計算されたとする。

最終的な合成標準不確かさの見積もり

定量値;Cs=0.254x100/0.5023=49.8　$\mu g\ g^{-1}$

$\mu g\ ml^{-1}$　　ml　　　g

故定量値の合成標準不確かさ(相対値)は以下の相対標準不確かさの合成で見積もることができる。

①秤量に関する相対標準不確かさ

$u_{C1,r}$ = 0.00012/0.5023=0.00024

②溶解後の定容操作に関する相対標準不確かさ

$u_{C2,r}$ = 0.077/100=0.00077

③標準液に関する相対標準不確かさ

$u_{C3,r}$ = 0.0052

④検量線法による測定に関する相対標準不確かさ

$u_{C4,r}$ = 0.0029/0.254＝0.0114

①②③④の4つの要因を合成して

$$u_{c,r} = \sqrt{0.00024^2 + 0.00077^2 + 0.0052^2 + 0.0114^2} = 0.013$$

u_C = 0.013/49.8=0.65 μg g^{-1}

拡張不確かさは、包含係数k=2(95%の信頼性)として

U= 0.65x 2=1.3 μg g^{-1}

以上より、試料中の金属元素Mの分析値として

49.8±1.3 μg g^{-1}

=50±1=(5.0±0.1)x10^{1} μg g^{-1}

と見積もることができた。RSDで表現すると2 %になる。

おわりに

- 分析をはじめ、あらゆる計測の際は、測定値の信頼性をいつも意識する。
- 正しい化学分析手順を確認してから、不確かさの見積もりを行う。
- 数値は一旦提示したら一人歩きする。数値がもたらす影響力を意識すること。
- 測定値を元に分析値を提示するのは分析者の責任である。

0011 0010 1010 1101 0001 0100 1011

michihisa.uemoto@meisei-u.ac.jp