



DG[2024-73]
バイオアナリシスに関連する統計学

はじめに

本基礎講座はJBF DG「バイオアナリシスに関連する統計学(2024-73)」のメンバーによって作成されたものです。

また、本基礎講座は、DGにて議論された内容を一部含みます。

そうした議論の内容は、メンバー個人およびDGの意見/考えであり、JBFとしての意見でないことをご理解ください。

DG[バイオアナリシスに関連する統計学]とは？

背景

- 統計学はバイオアナリシスのデータの信頼性評価などに大変重要
- データ分析や応用数学といった分析化学や生物化学の知識とは異なる素養が求められるため、修学や理解のハードルが少々高い

ガイドラインに従って機械的に統計手法を使用するのではなく、バイオアナリシスに関連する統計用語・統計解析を基礎から理解し、本質的な議論をすることを目的に組織されたDG

所属メンバー

氏名	所属組織	役割
栗栖 泰之介	ファイザーR&D合同会社	DGリーダー
河野 憲史	株式会社東レリサーチセンター	
橘田 久美子	シミックファーマサイエンス株式会社	
田中 奈那	株式会社サンプラネット	
野田 和揮	大塚製薬株式会社	
浜田 梨沙	株式会社住化分析センター	
山田 翔太	シミックファーマサイエンス株式会社	
大和 遼	メディフォード株式会社	
中村 将俊	ファイザーR&D合同会社	オブザーバー

※本ポスターに関するご質問はメールでHironosuke.Kurisu@pfizer.comにお願いします。



バイオアナリシスに おける統計学の重要性

なぜバイオアナリシスで統計学を利用するのか？

⇒データの誤差やばらつきを適切に評価し、

信頼性の高い結論を導き出すため

1. データのばらつきを評価する

個体差 等に起因するばらつきが大きい生物学的データを定量的に評価する

2. 誤差を管理する

誤差の影響を最小限に抑える方法を提供し、結果の精度を高める

3. データを比較する

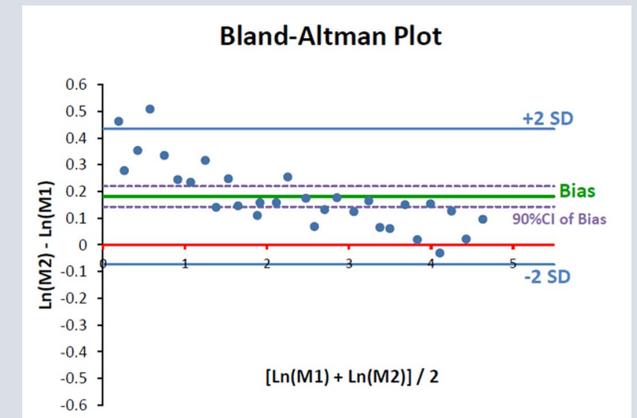
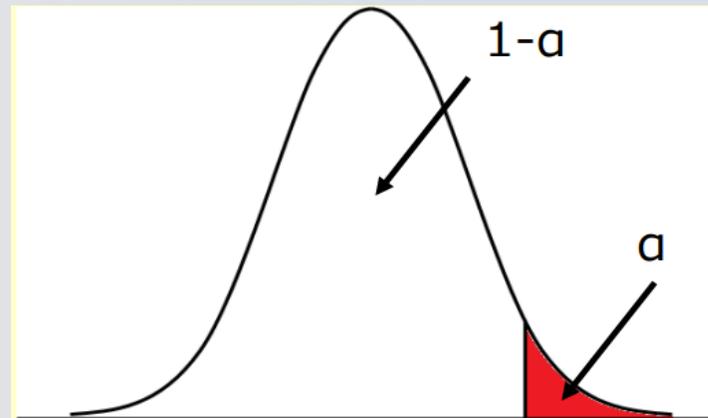
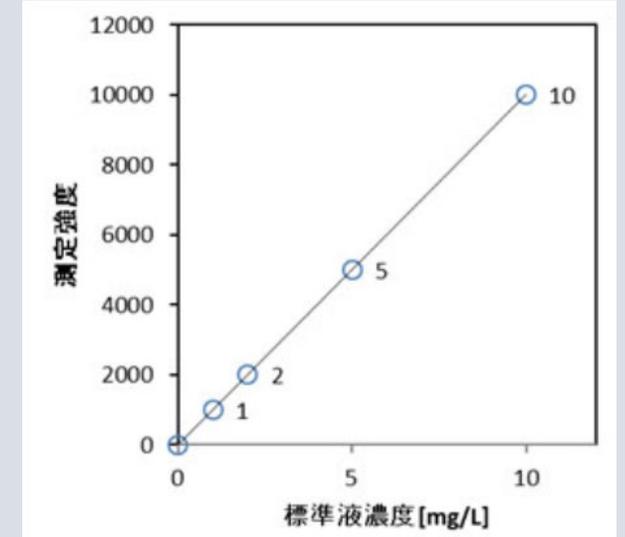
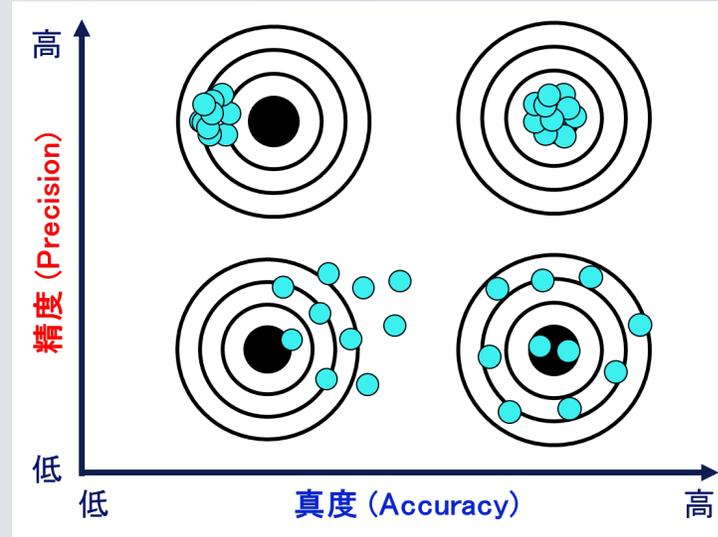
異なる条件やグループ間でのデータを比較し、複数群の違いを明確化する

バイオアナリシスで活用する統計解析の一例

バイオアナリシスにおいては、

- データの信頼性評価
- 重みづけ補正した検量線作成
- ADA分析のカットポイント設定
- クロスバリデーション時の結果の解析

など、様々な状況で
統計解析が使用される。





真度(Accuracy)と 精度(Precision)

真度と精度 ～その具体的な意味と意義①～

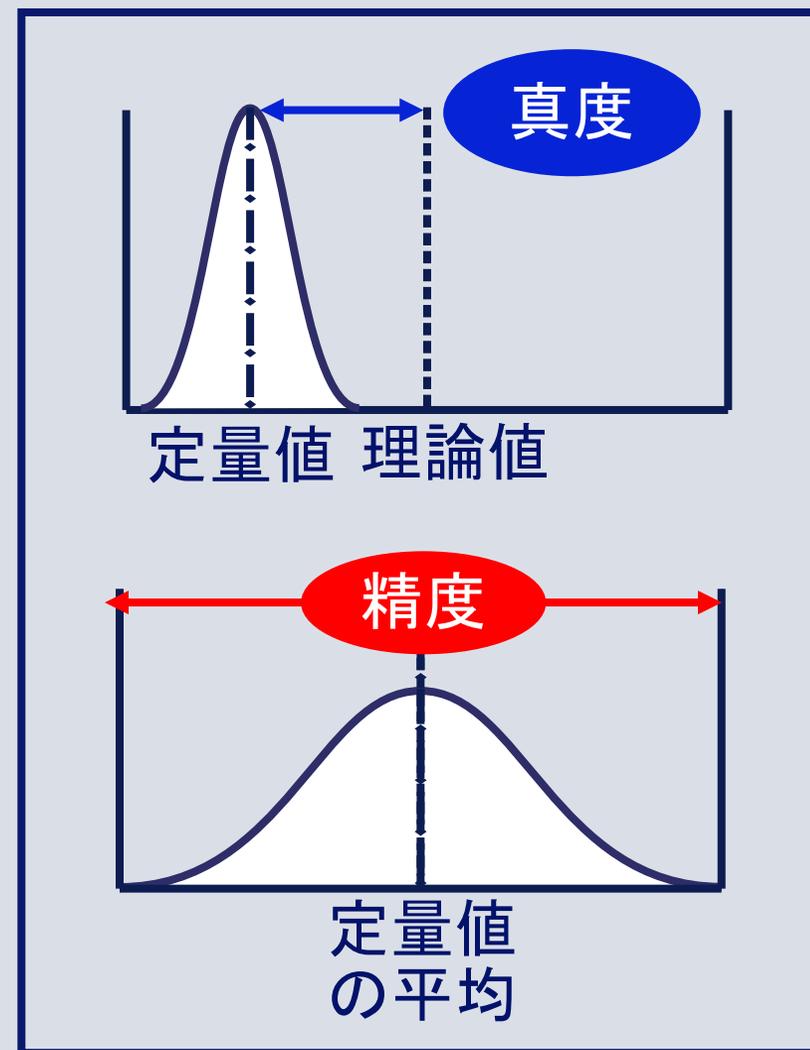
真度 (Accuracy)

真度は、分析対象物質の
定量値と理論値との一致の程度のこと。

精度 (Precision)

精度は、繰り返し分析によって
得られる定量値のバラツキの程度のこと。

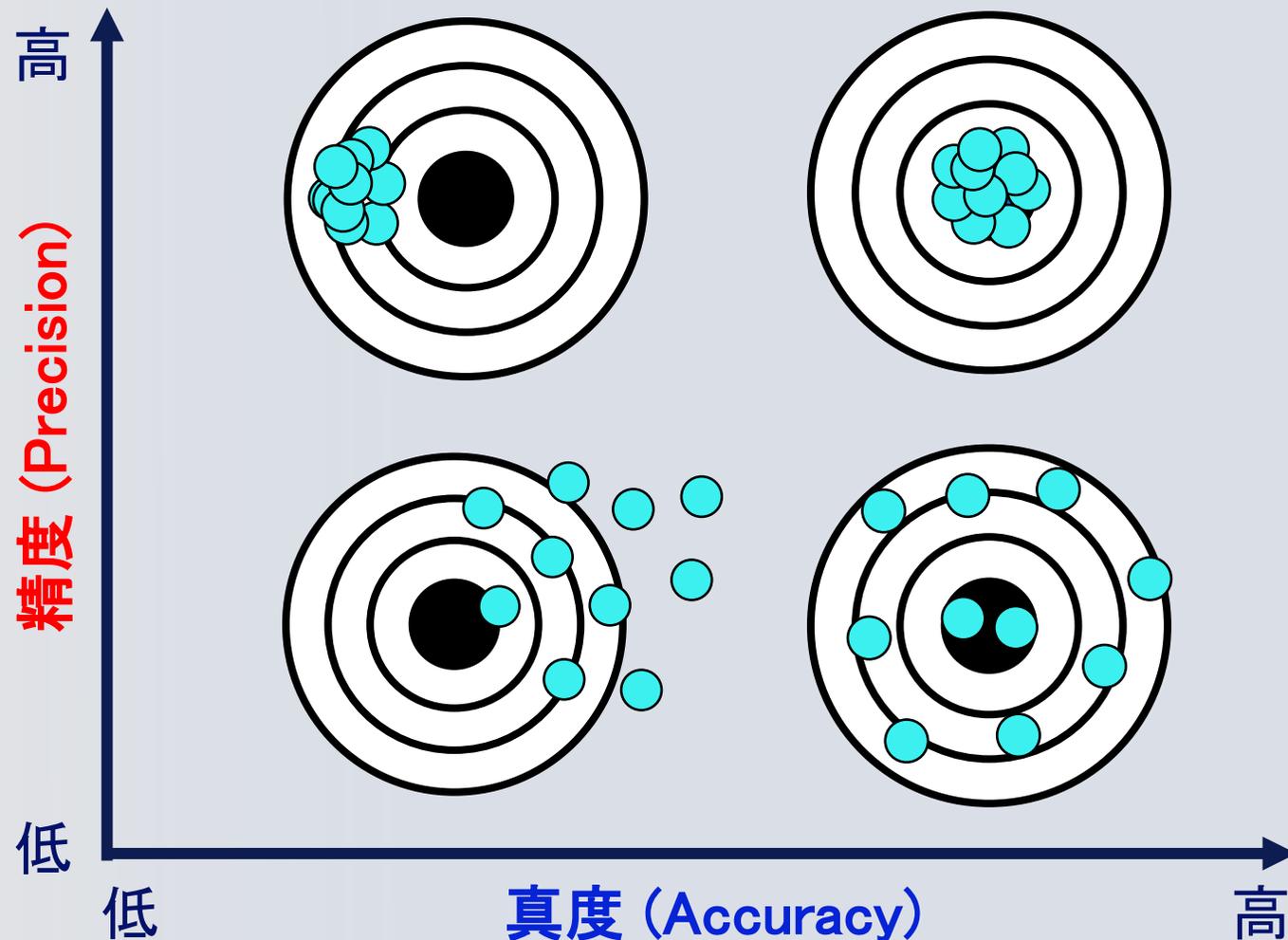
同様の測定を実施しても、様々な要因の
誤差が生じることにより定量値は毎回異なる
真度と精度を用いれば、誤差の性質を正しく理解できる



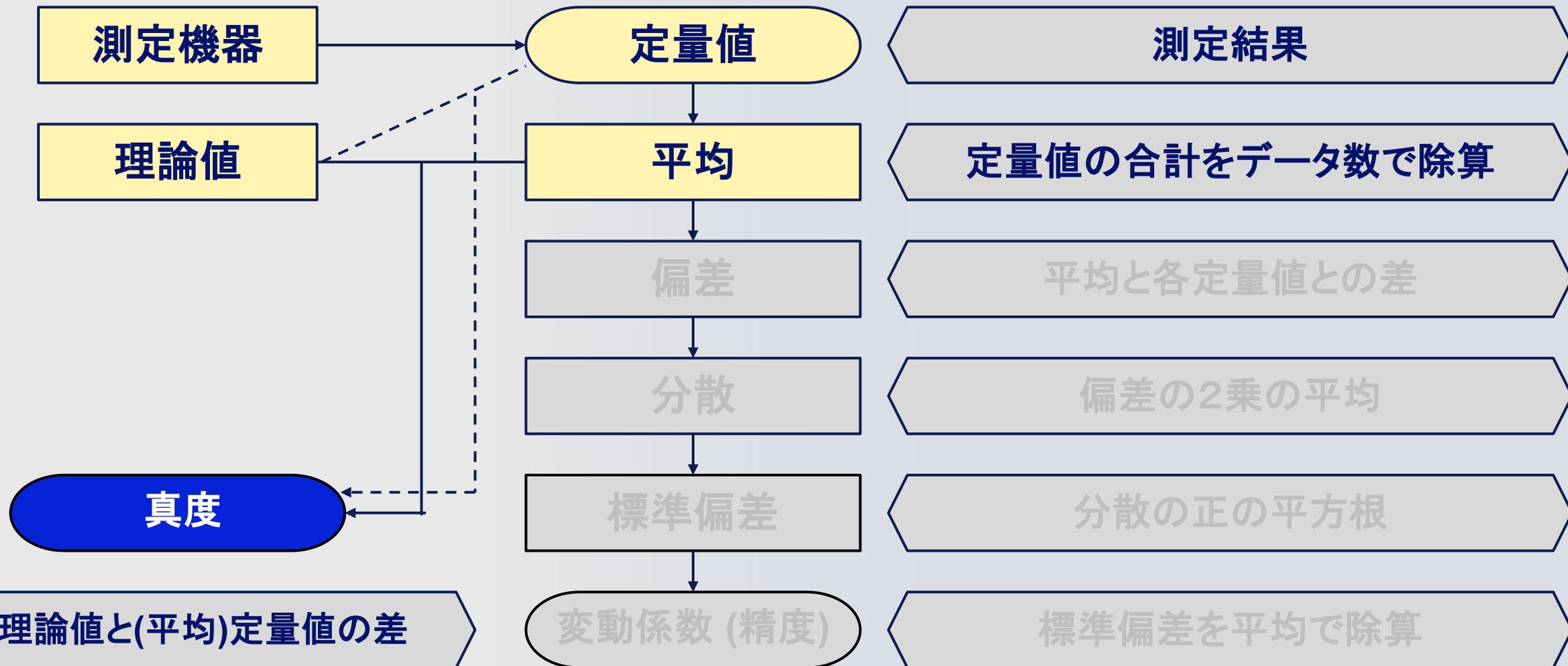
真度と精度 ~その具体的な意味と意義②~

真度と精度の関係は
的の図で解釈することができる

- 真度と精度が両方高い場合、
中心に点が集中する(右上)
- 真度と精度が両方低い場合、
点は中心から外れながら散る(左下)



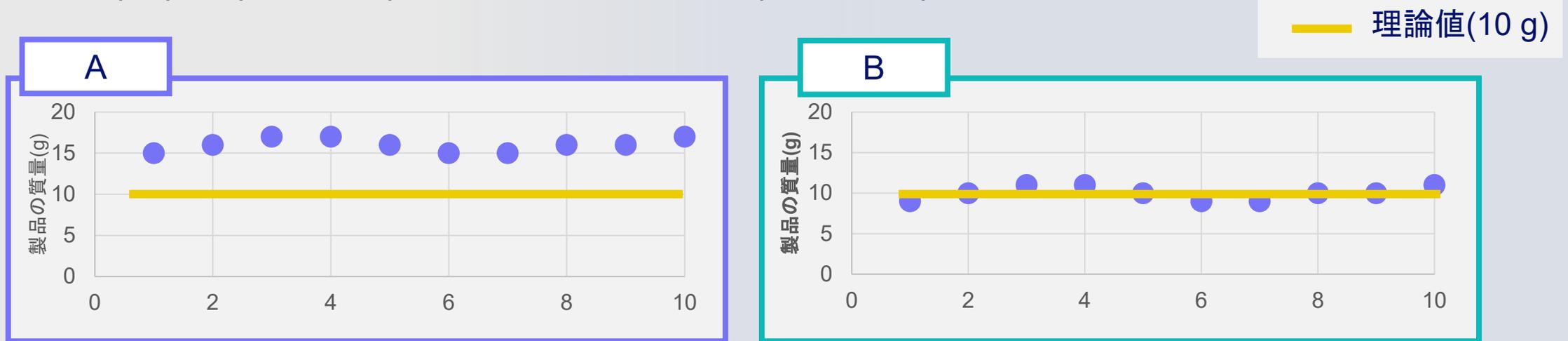
真度を求める際のフロー



真度 (Accuracy)

真度は、定量値あるいはその平均値が理論値にどの程度近いかを示す値である。

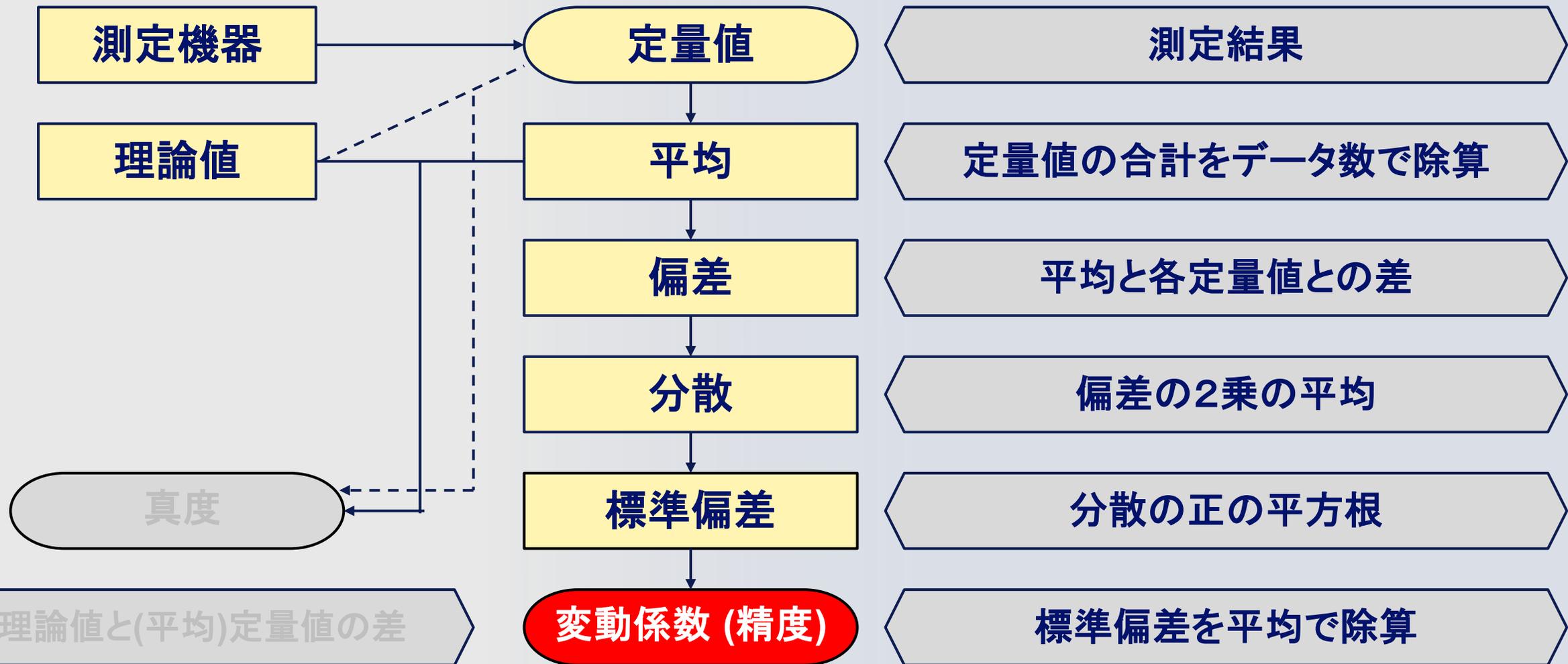
真度(%) = (定量値(or 定量値の平均値)/理論値) x 100で表される。



	1回目	2回目	3回目	4回目	5回目	6回目	7回目	8回目	9回目	10回目	平均	真度(%)
A	15	16	17	17	16	15	15	16	16	17	16	160
B	9	10	11	11	10	9	9	10	10	11	10	100

AとBのデータセットを比較したとき、より真度が高いのはB

精度を求める際のフロー



平均 (Mean)

平均とは、データの中心的な傾向を示す指標である。

算術平均(相加平均)

各データを足し算してデータ数で割った値。最も身近な「平均値」

$$\frac{a_1 + a_2 + \dots + a_n}{n}$$

幾何平均(相乗平均)

各データを掛け算した結果のn乗根をとった値。C_{max}やAUCなどのデータの平均を算出する時に使用する。

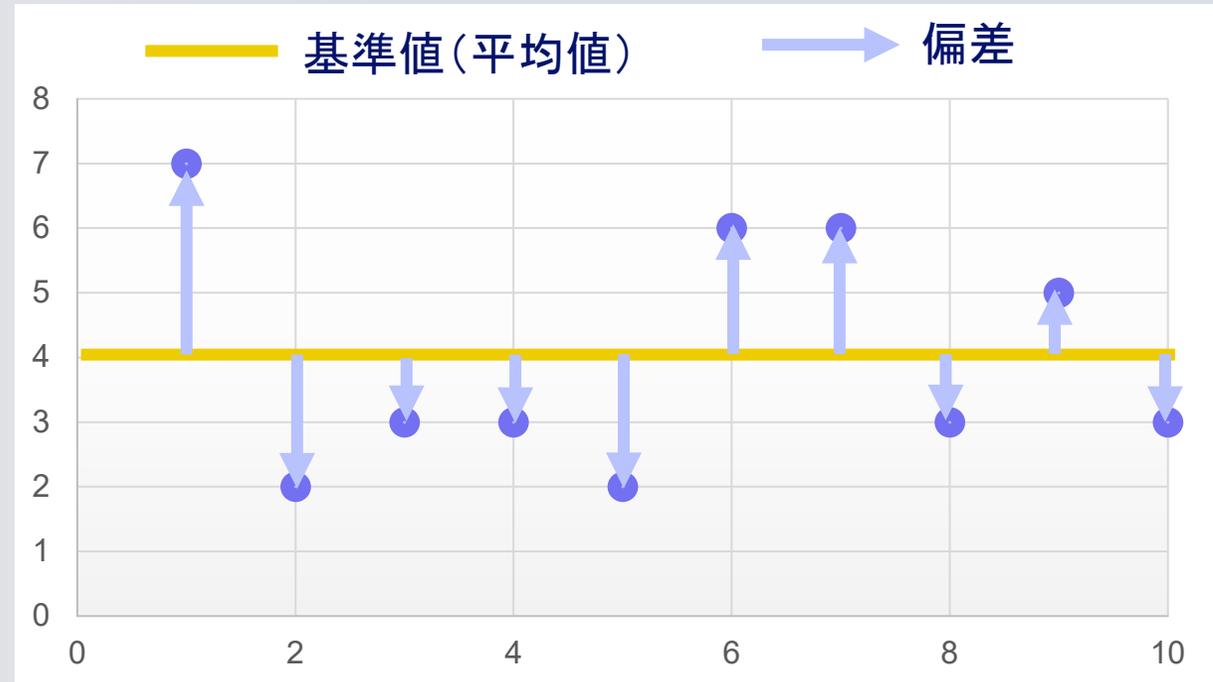
$$\sqrt[n]{a_1 \times a_2 \times \dots \times a_n}$$

データの分布やデータの使用目的に従って、
適切な平均値を選択することが重要

偏差 (Deviation)

偏差は、基準とする値と各データとの差である。

- それぞれの値が基準値からどの程度離れているかを表す
- 偏差の算術平均は0になる



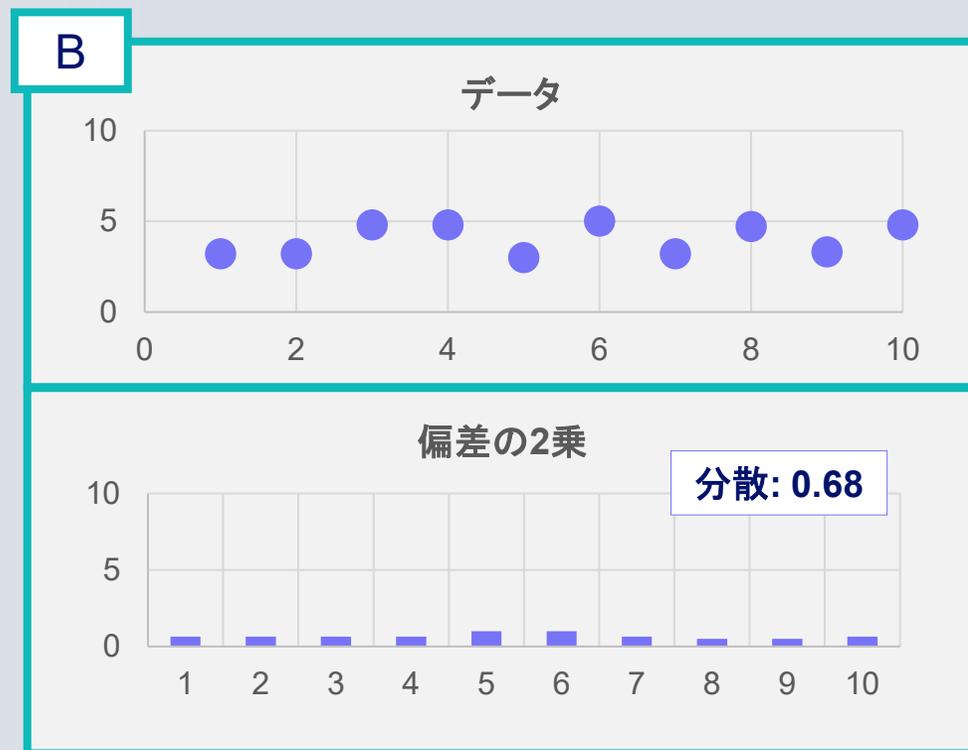
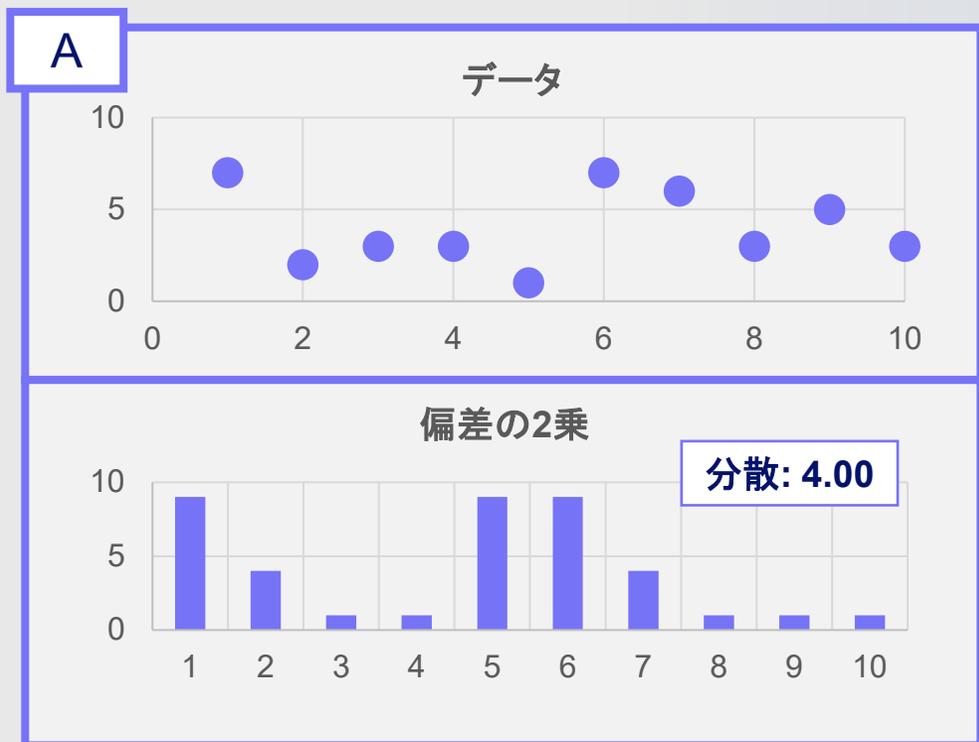
データ	7	2	3	3	2	6	6	3	5	3
平均値からの偏差	+3	-2	-1	-1	-2	+2	+2	-1	+1	-1

偏差の算術平均は0になってしまうため、

偏差を比較してもどちらがバラつきが大きいかを判断することは困難

分散 (Variance)

分散は、データがどの程度ばらついているかを表す指標である。分散が大きい程、データ全体のばらつきも大きい。偏差の2乗の算術平均で表される。



A, Bどちらも平均値は4だが、Bの方が値のばらつきが少ない(分散が小さい)

標準偏差 (Standard Deviation, SD)

標準偏差は、データのばらつきを表す数値であり、分散と異なり平均値と同単位であるため、平均値を基準としたばらつきを表現できる。

- 分散の平方根をとることで求めることができる
- 標準偏差が小さい程、データは平均値に集中し、一方、大きい程ばらつきが大きいことを示す

$$SD = \sigma = \sqrt{\frac{(a_1 - \bar{a})^2 + (a_2 - \bar{a})^2 + \dots + (a_n - \bar{a})^2}{n}} = \sqrt{v^2}$$

値	単位	
平均値	g	ng/mL
分散	g ²	ng ² /mL ²
標準偏差	g	ng/mL

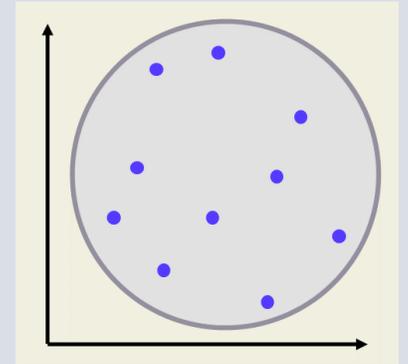
例えば、平均値の単位がgである場合、分散の単位はg²、標準偏差の単位はgとなる

標準偏差にも種類がある！

①母集団の標準偏差

母集団の標準偏差そのもの

$$= \sqrt{\frac{(a_1 - \bar{a})^2 + (a_2 - \bar{a})^2 + \dots + (a_n - \bar{a})^2}{n}}$$



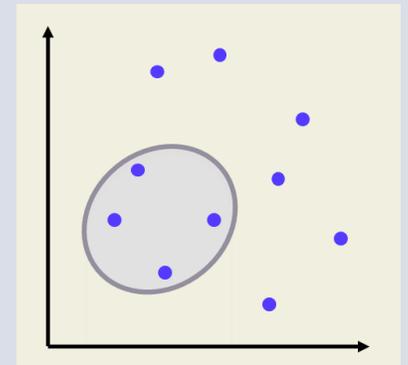
「母集団全体」の
データから算出

②標本の標準偏差

(不偏標準偏差)

母集団の標準偏差の推定値

$$= \sqrt{\frac{(a_1 - \bar{a})^2 + (a_2 - \bar{a})^2 + \dots + (a_n - \bar{a})^2}{n - 1}}$$



「母集団の一部」の
データを抽出して算出

変動係数 (Coefficient of Variation, CV) = 精度 (precision)

変動係数は、データのばらつきを表す数値であり、

標準偏差 ÷ 平均値

で算出する。

- 平均値に対する相対的なばらつきを確認できる
- 一般的にはその百分率を**精度**と呼ぶ

サンプル	平均値	標準偏差		変動係数	精度
● × 小学校の1年生の体重	25.0 kg	8.0 kg	⇒	0.32	32.0 %
△ □ (株) の男子社員の身長	170.0 cm	8.5 cm	⇒	0.05	5.0 %

単位や平均値が異なるサンプルのばらつきを比較できる

乖離度 (the percent difference, %difference)

乖離度は、初回の定量値と再分析による定量値との相違の程度であり、両者の平均に対する両者の差をパーセント表記した値である。

$$\text{乖離度}(\%) = \frac{\text{再分析の定量値} - \text{初回の定量値}}{\text{両者の平均値}} \times 100$$

- 初回の定量値と再分析値の一致度を確認できる
- 分析法が一貫して正確な結果を提供するかを評価する
- 全体の分析プロセスが適切に機能しているかを評価する

サンプル	初回測定 定量値	再分析 定量値	乖離度
A	2.00	2.25	11.8 %
B	2.00	1.75	-13.3 %
C	2.00	3.00	40.0 %



正規分布と外れ値

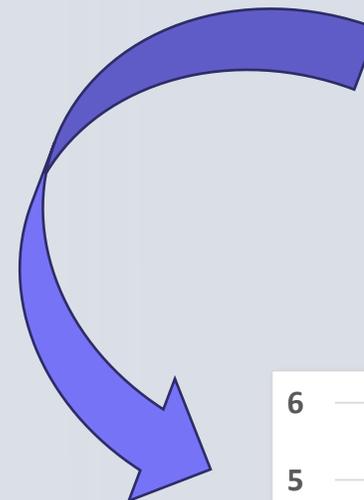
ヒストグラム (Histogram)

ヒストグラムは、度数分布表から作成したグラフである。

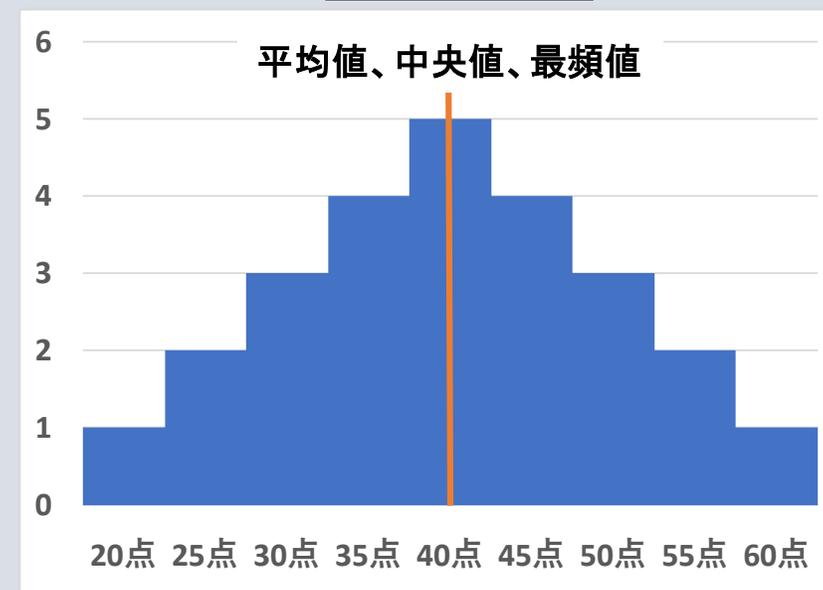
- データの散らばり具合や性質を表すのに適している。
 - 横に広がっていればバラつきが大きく、シャープであればバラつきが小さい
 - 左右対称であれば正規性が高く、片方に偏っていれば、正規性が低い
- 横軸は階級、縦軸は度数を示す。
 - 階級: 度数を集計する区間
 - 度数: 各階級のデータ数

度数分布表

階級	階級値	度数
18点~22点	20点	1
23点~27点	25点	2
28点~32点	30点	3
33点~37点	35点	4
38点~42点	40点	5
43点~47点	45点	4
48点~52点	50点	3
53点~57点	55点	2
58点~62点	60点	1



ヒストグラム



正規分布 (Normal Distribution)

正規分布は、統計学における最も重要な確率分布の1つであり、横軸を確率変数、縦軸を確率密度にとると、

★左右対称

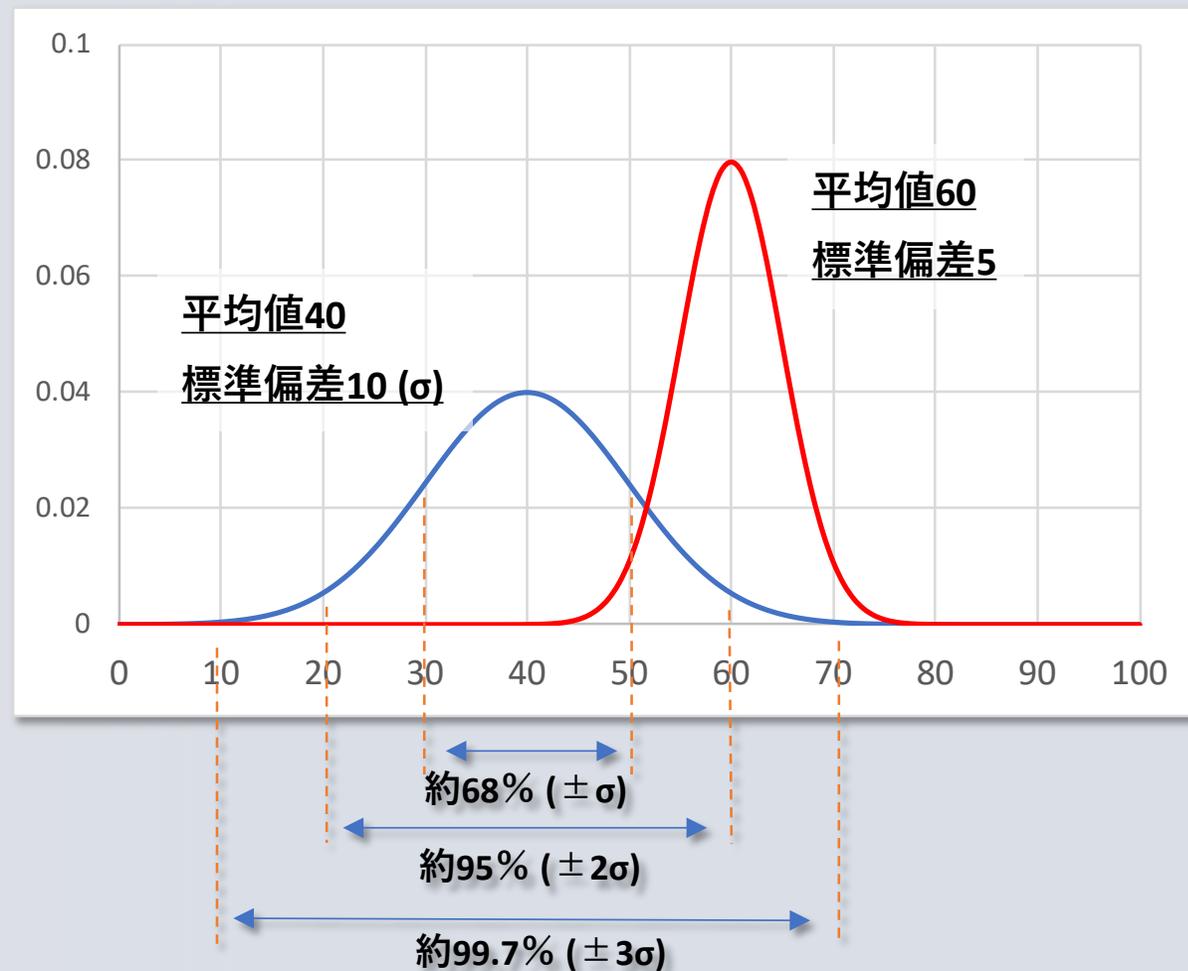
★平均値 = 中央値 = 最頻値

となる特徴を持つ

- データが正規分布に従うことを **正規性を持つ**と表現する
- 実験誤差など、多くのデータは正規分布に従う(正規性を持つ)と仮定される
- 上記の仮定に基づいて、様々な解析を行うことがある

中央値: データを小さい順(大きい順)に並べたとき、中央に位置する値

最頻値: データの中で最も頻繁に出現する値



なぜ正規分布が重要なのか？

通常、分析結果のバラつきや誤差を完全に除去することは不可能であるが、それらは正規分布に従って発生することが知られている。

「分析結果や誤差などの様々な事象は正規分布に従う」という前提に基づけば、種々の現象を正規分布というモデルを介して、より深く理解できるようになる。

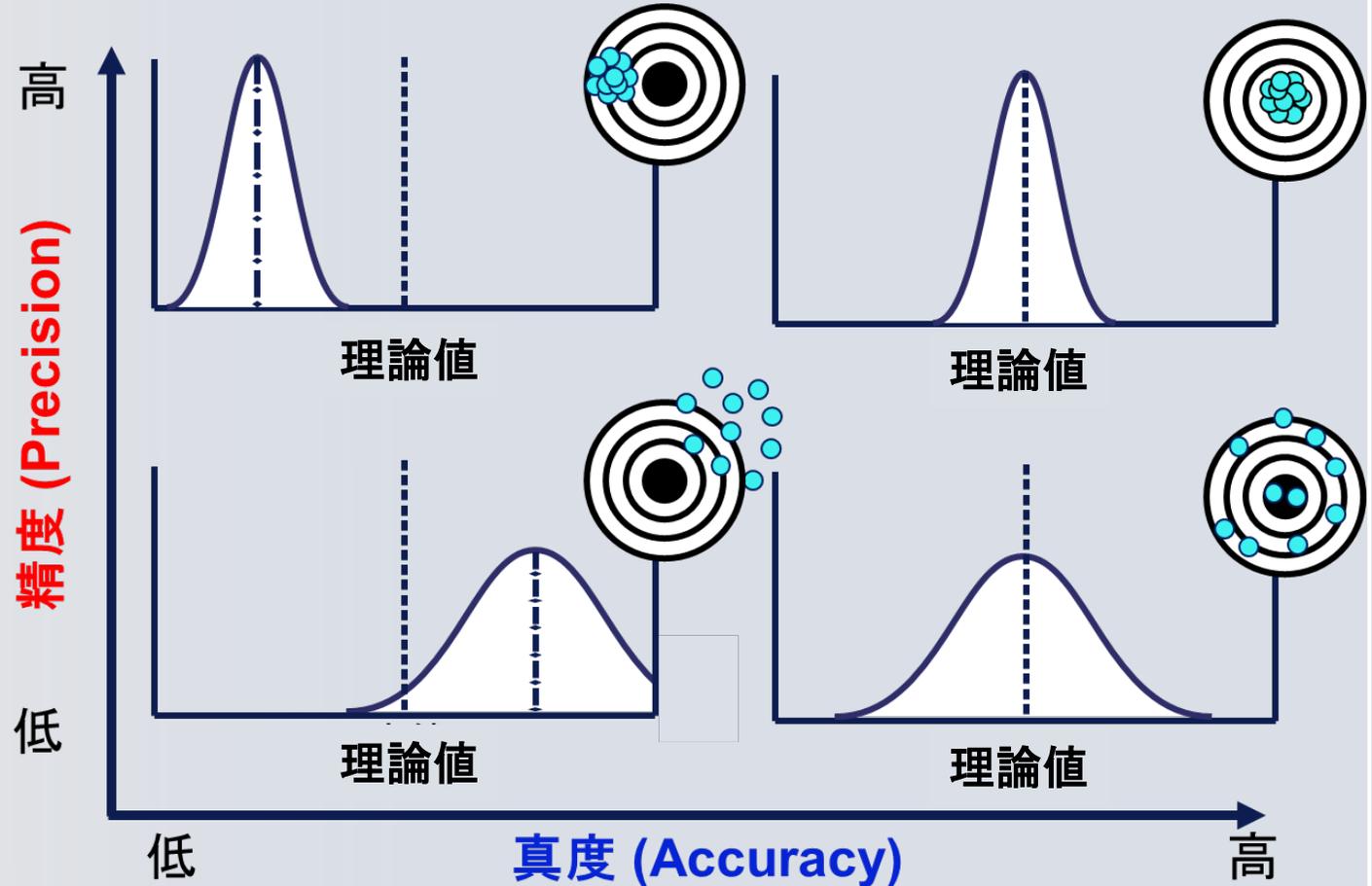
例えばP10で説明した

真度 (理論値との離れ具合)

精度 (バラつき具合)

的的の図は、

正規分布を加えると右のように表現できる

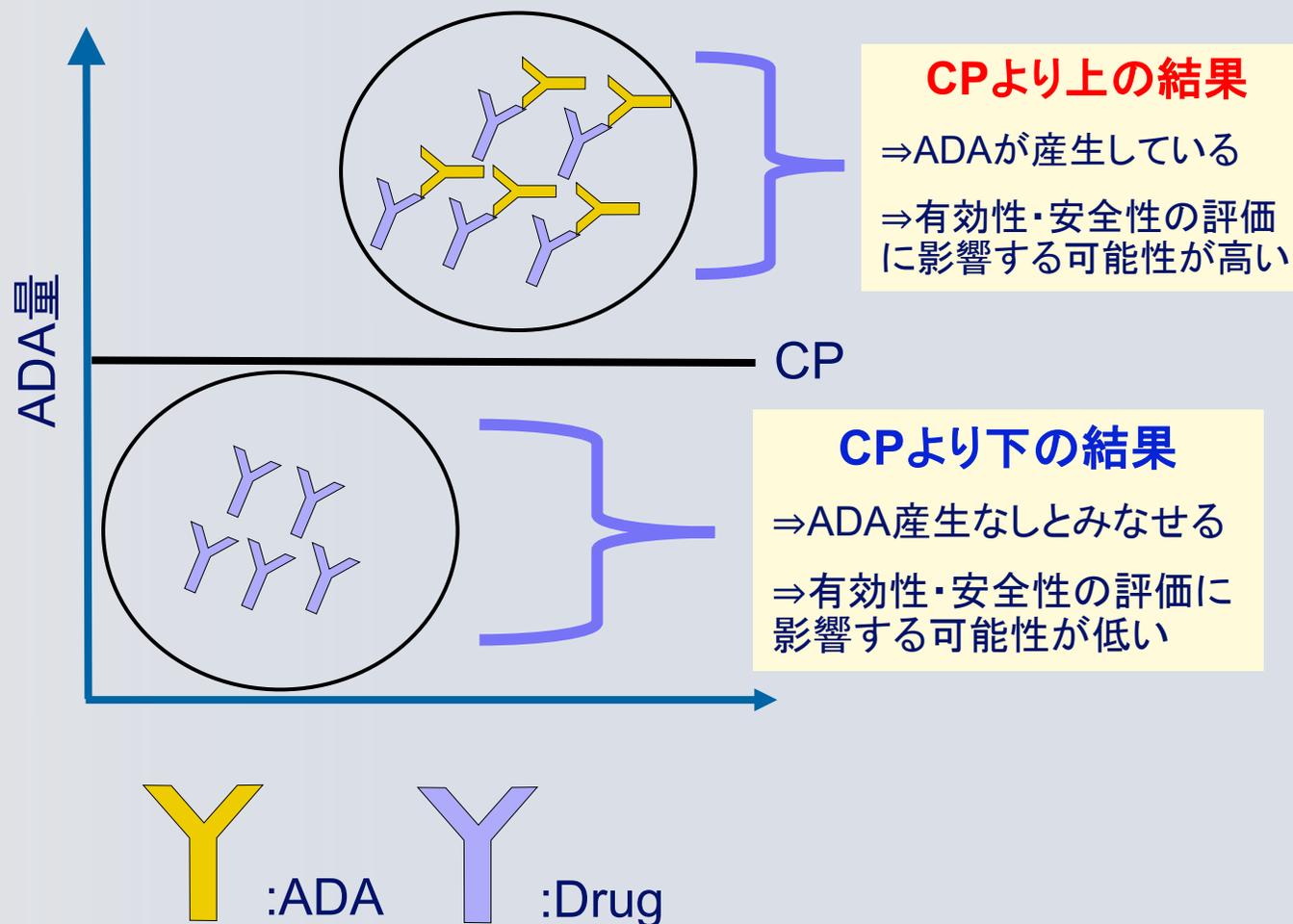


正規分布活用の一例: ADA分析のCP設定

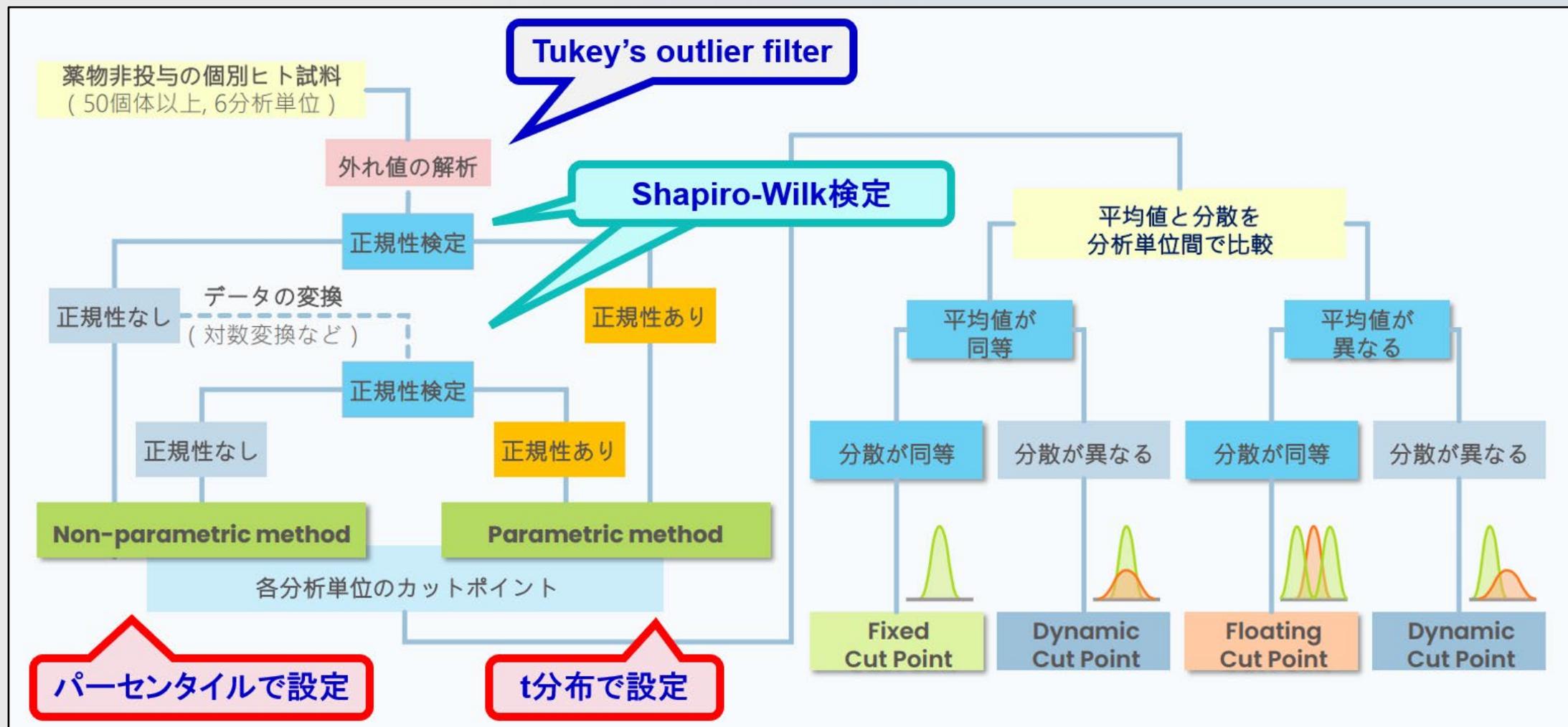
バイオアナリシスでの正規分布の活用方法としては、例えば抗薬物抗体分析(ADA分析)のCP(カットポイント)設定が挙げられる。

⇒予め設定したCPより上か下か、によってADA産生の有無を判定する

※ADAの産生の有無と産生量、薬物の安全性・有効性にもたらす影響を評価することは極めて重要である。



ADA分析のCP設定のフローチャート



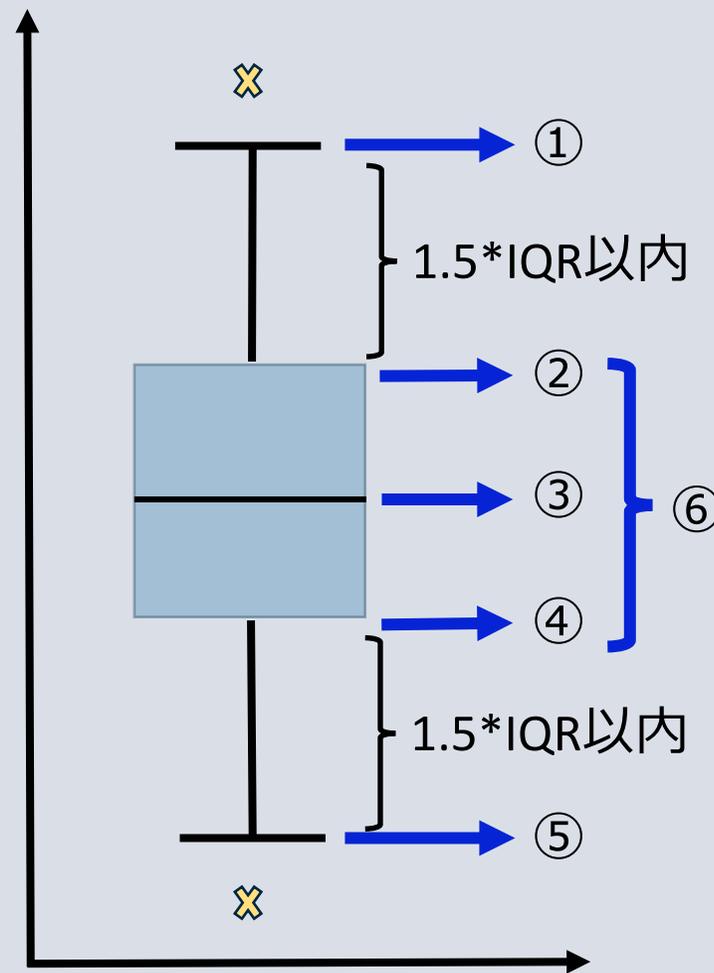
箱ひげ図 (Box Plot)

箱とひげを合わせたグラフを「箱ひげ図(Box-plot)」と呼ぶ。

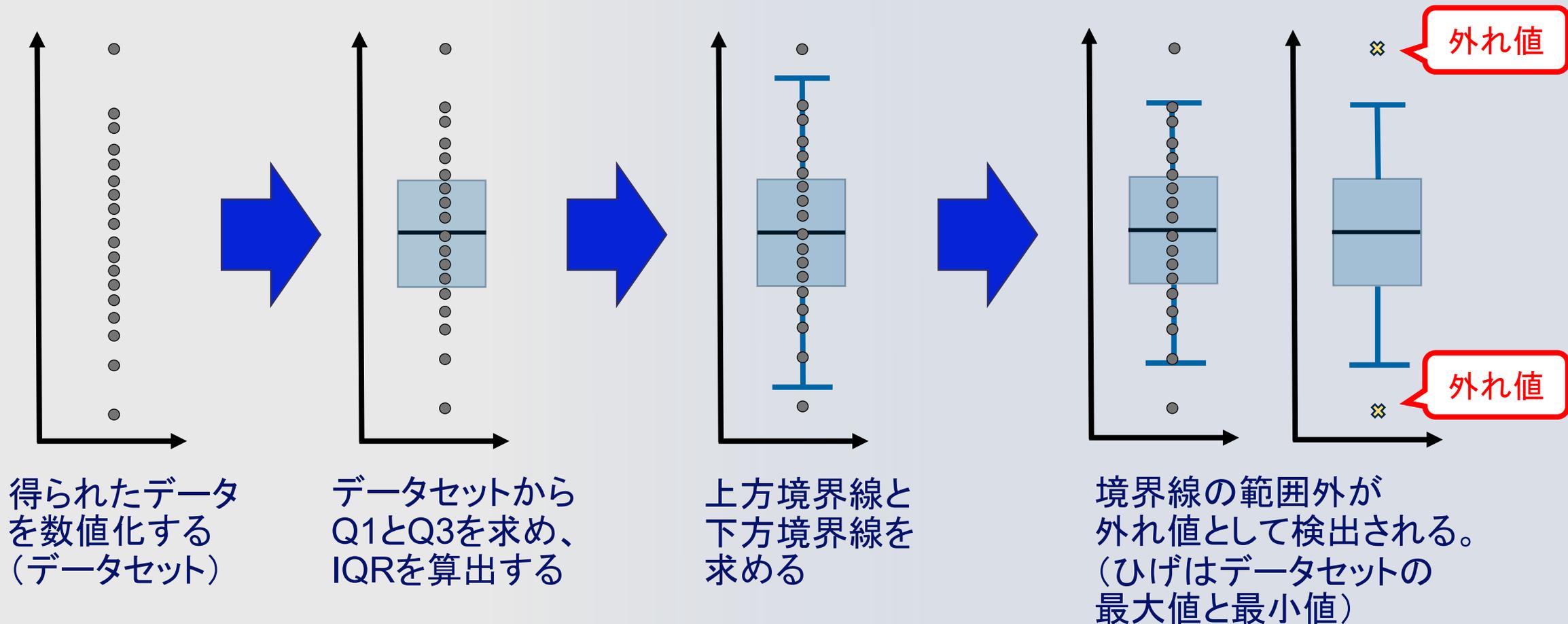
箱: データの下位25%~上位25%を含む領域(②~④まで)

ひげ: 箱のQ1とQ3から1.5*IQR離れた境界線に収まる、
データセットの最大値もしくは最小値(①~②と④~⑤)

- | | |
|---------------------------|---------------------------|
| ①上方境界線 | $Q3 + 1.5 \cdot IQR$ までの値 |
| ②第3四分位数(Q3) | データの上位25%の境界の値 |
| ③第2四分位数(Q2) = 中央値(Median) | |
| ④第1四分位数(Q1) | データの下位25%の境界の値 |
| ⑤下方境界線 | $Q1 - 1.5 \cdot IQR$ までの値 |
| ⑥四分位範囲(IQR) | 箱にあたる部分で、データの50%が含まれる範囲 |



外れ値解析 ~Tukey's outlier filter~



「対象となるデータセットが正規分布に従う」

という仮定に基づく解析であることに注意が必要

データセットの正規性を確認する①

統計的仮説検定による方法

例) Shapiro-Wilk検定

p値が0.05以上で正規性を持つ、と認める

$$W = \frac{\left(\sum_{i=1}^n a_i x_{(i)}\right)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

統計ソフトで解析する

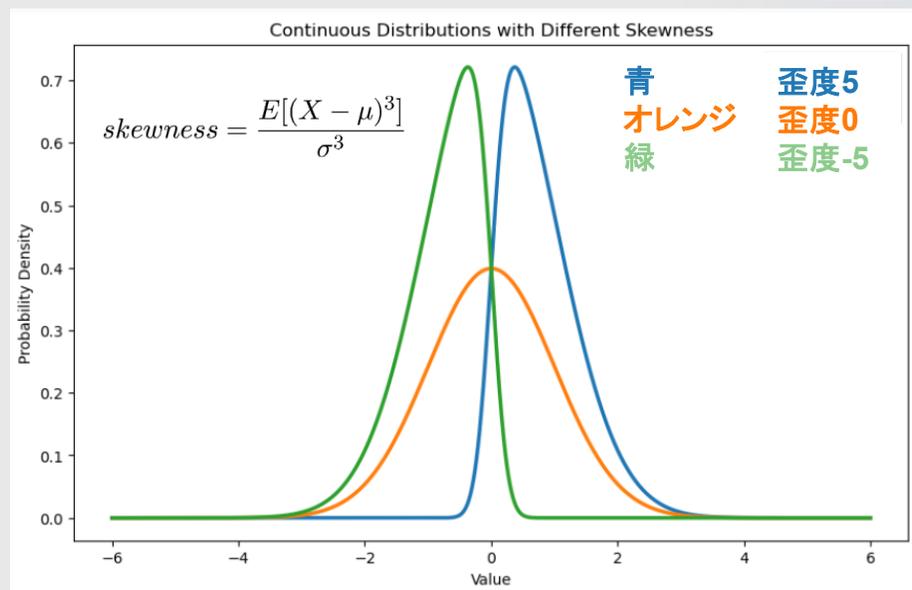
統計的仮説検定(Statistical Hypothesis Testing)とは？

母集団に関するある仮説が、予め定めた有意水準を基準として、
統計学的に成り立つか否かを標本のデータを用いて判断すること。
以下の手順で実施され、結論を導くためには背理法が使用される。

1) 仮説の設定 ⇒ 2) 有意水準の決定 ⇒ 3) 検証 ⇒ 4) 結論

データセットの正規性を確認する②

歪度と尖度を利用して、正規性を評価することがある。

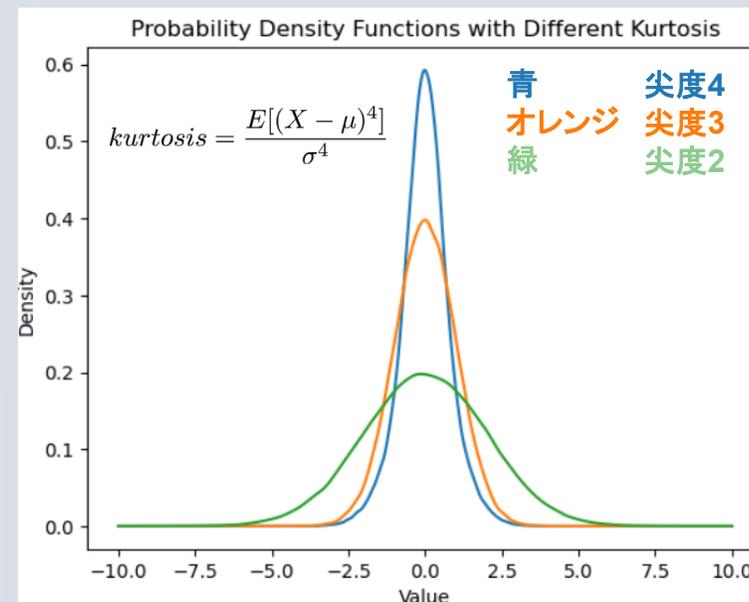


歪度 (Skewness)

意味: データの非対称性を示す

正の歪度: 右に長い尾を持つ分布 (右に歪んでいる)

負の歪度: 左に長い尾を持つ分布 (左に歪んでいる)



尖度 (Kurtosis)

意味: データのピークの鋭さを示す

高い尖度: 鋭いピークを持つ分布 (尖っている)

低い尖度: 平らなピークを持つ分布 (平坦)



検量線と重みづけ

濃度未知の検体の濃度を確認するプロセス

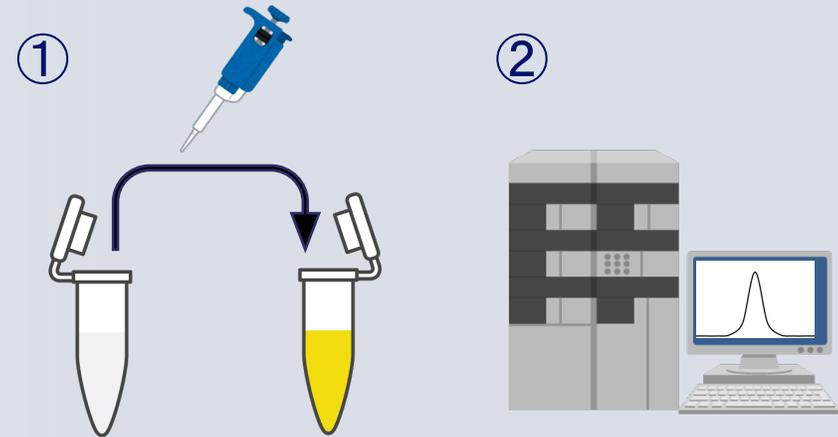
①分析対象物質を含む標準溶液をブランクの生体マトリックス等に添加し，既知濃度の検量線用標準試料を調製

②前処理後，HPLC，LC-MS/MS等で測定

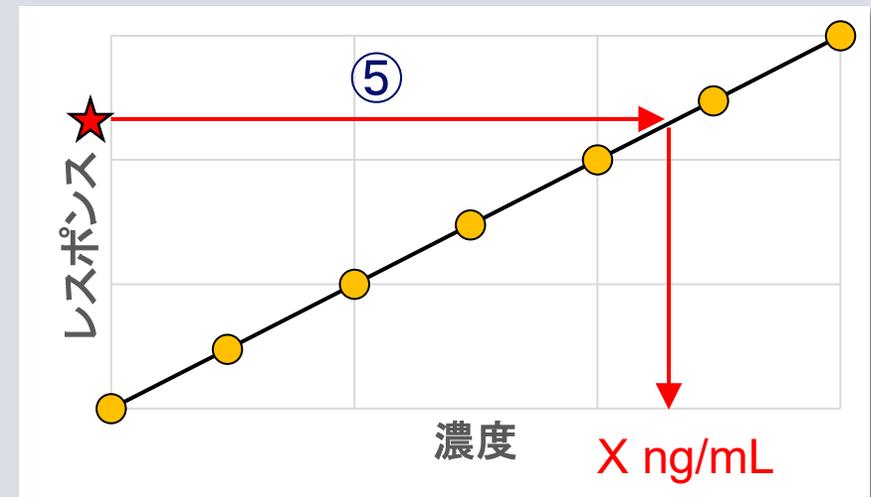
③x軸に濃度，y軸にレスポンス(ピーク面積，面積比等)をプロット(●)し，散布図作成

④回帰分析を実施し，標準曲線(検量線)作成

⑤未知試料を測定し，レスポンス(★)を標準曲線にあてはめ，未知検体の濃度を算出



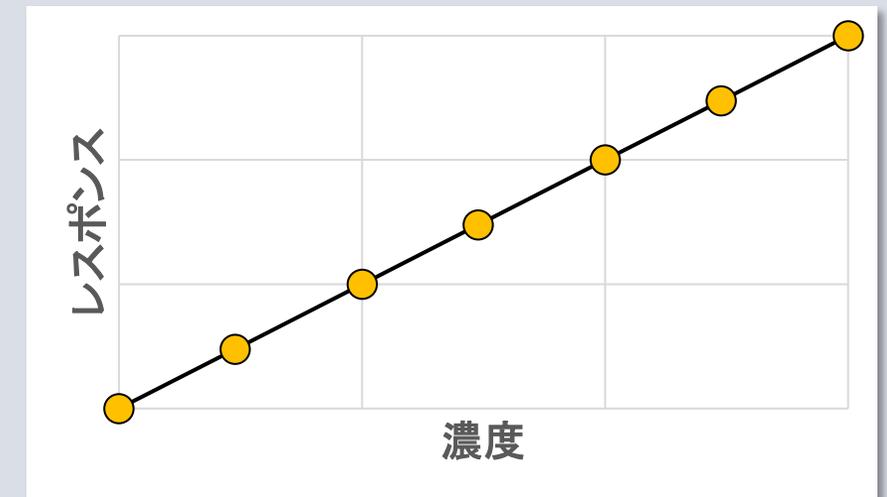
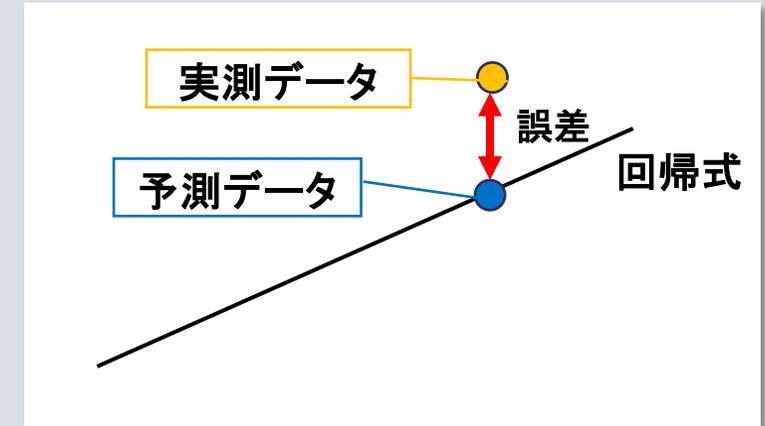
③散布図 & ④回帰分析



回帰分析 ～特に線形回帰について～

回帰分析は、X軸(濃度)とy軸(レスポンス)をモデル化すること。バイオアナリシスでは回帰分析で得た回帰式を標準曲線(検量線)として利用する

- 係数が1つの場合は単回帰 (LC-MS/MS)
 複数がある場合は重回帰 (LBA)
- すべてのデータで、実測データと回帰式上の予測データの誤差が最も小さくなる回帰式が理想
- 散布図に回帰式を重ねることで、回帰式が適切か、濃度によって誤差に差があるか(重みづけが必要か)等を判断できる



LC-MS/MSの検量線(回帰直線)の求め方

回帰直線の一般式: $y = ax + b$

実験で得られた観測値 (y_i) と、回帰直線から計算される予測値 (\hat{y}) との差 ($e_i = \hat{y} - y_i$) の二乗和が最小になるように係数 a 、 b を決める。

※ e_i には正負両方があるので、二乗和にする必要がある

EXCELで計算する場合は、

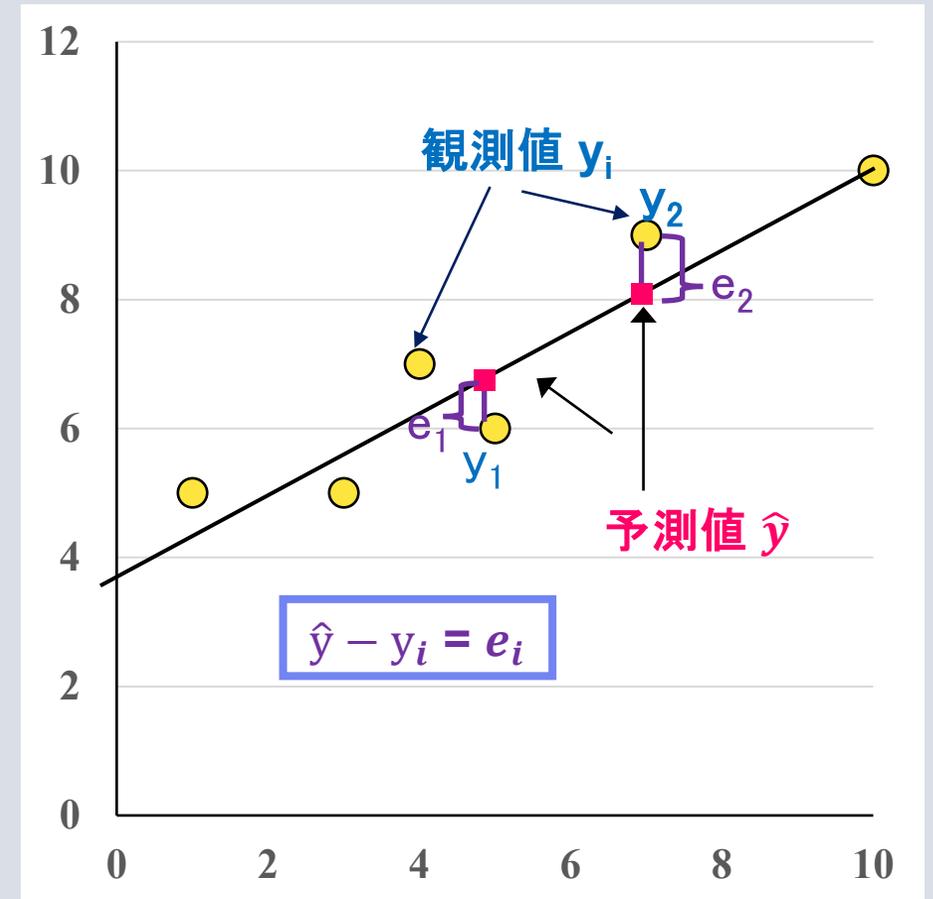
- ・傾き (a) = SLOPE (既知の y , 既知の x)
- ・切片 (b) = INTERCEPT (既知の y , 既知の x)

あるいは、

グラフ上で右クリック後、

[近似曲線の追加⇒線形近似⇒グラフに数式を表示する]

で算出可能



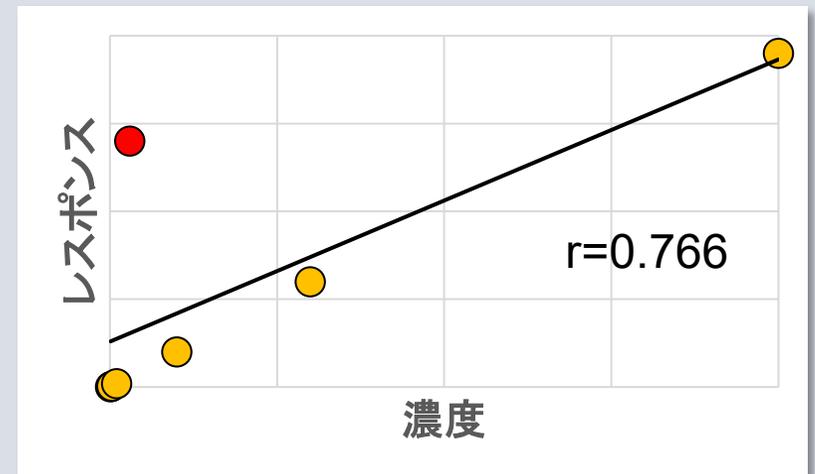
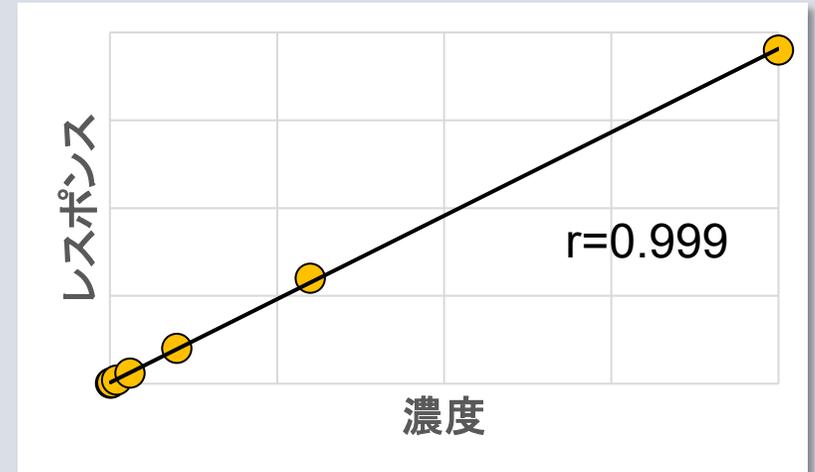
※ \hat{y} (y hat) は回帰分析でモデルが予測した y の値を示す。

相関係数 (correlation coefficient, r)

相関係数は、x軸とy軸の線形関係の強さと方向を示す指標である。主にピアソンの積率相関係数が使用される

- 相関係数 = $\frac{xとyの共分散}{(xの標準偏差) \times (yの標準偏差)}$
- 1の場合は正の相関, -1の場合は負の相関
(-1から1の範囲を取る)
- 相関係数の算出後は散布図と照らし合わせ、結果が外れ値の影響にされているかどうかを確認することが重要

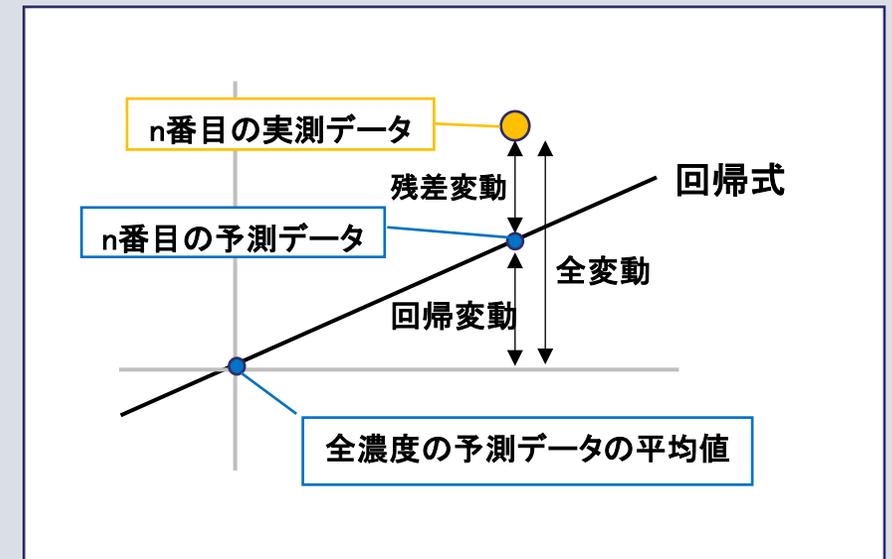
共分散: 2つの変数がどの程度一緒に変動するかを示す統計量



決定係数 (coefficient of determination, R^2)

決定係数は、回帰式の説明力を表す指標である。

- 決定係数 = $\frac{\text{回帰変動の平方和}}{\text{全変動の平方和}}$ (= 相関係数²)
- $0 \leq R^2 \leq 1$ の範囲を取る
- 1に近い程、回帰式が実態をよく説明している
- 決定係数が高いとは、「残差変動が全変動に対してどれだけ少ないか」を表す
- 相関係数の2乗であるため、必ず正の値となる



全変動: 実測データと平均値の差

回帰変動: 予測データと平均値の差

残差変動: 実測データと予測データの差

重みづけ (Weighting)

検量線の基本は線形回帰モデル式で算出される

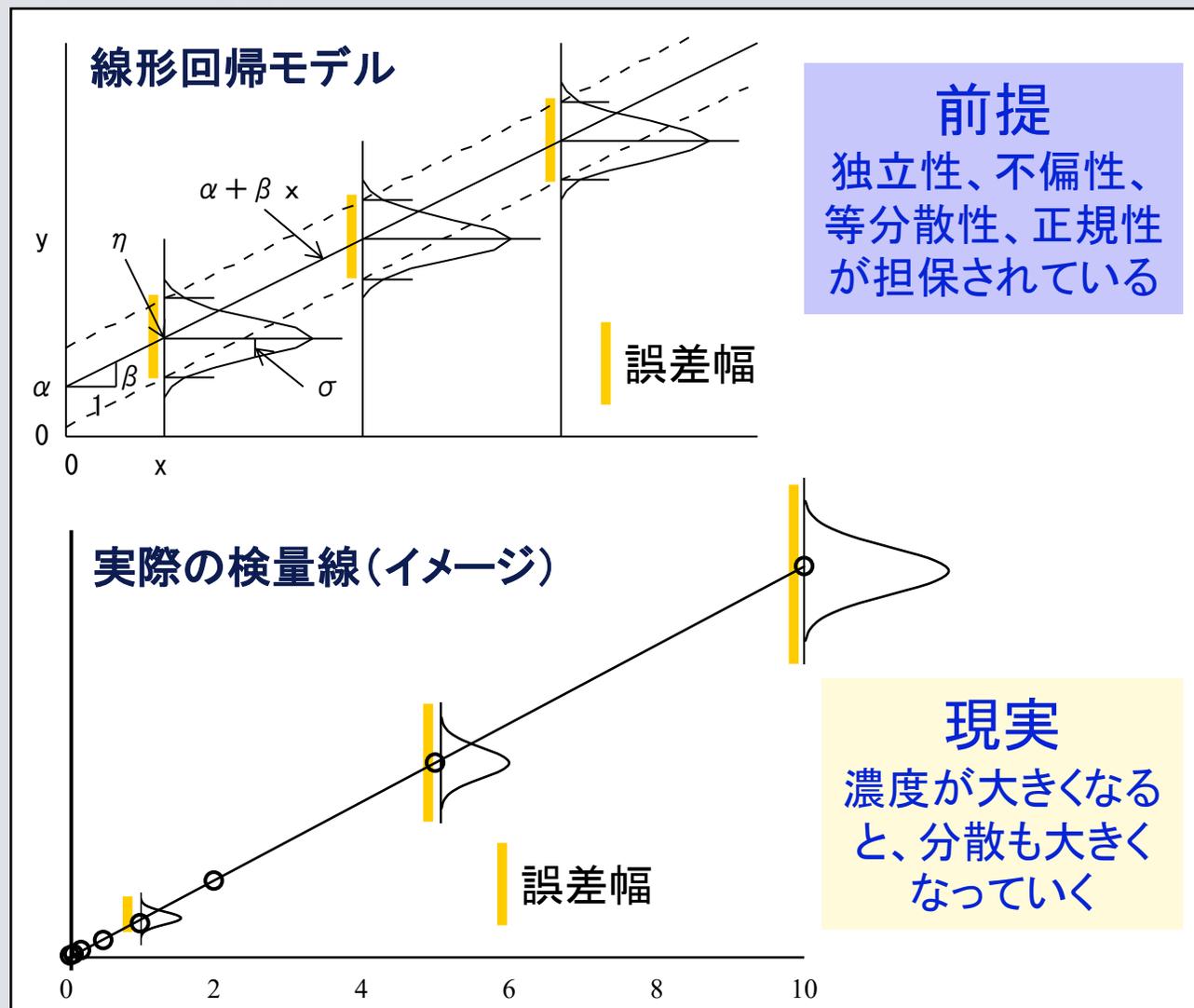
$$\text{モデル式 } y = a + bx = \alpha + \beta x + \varepsilon$$

線形回帰モデルの前提は、

[独立性、不偏性、等分散性、正規性]

実際の検量線は、濃度範囲が広く、濃度に比例して誤差が無視できない程大きくなるので、等分散性が成立しなくなり、高濃度域の誤差が低濃度域に大きく影響する。

⇒ 各濃度域での誤差を「等分散」に近づけるように補正する方法を重みづけという。



重みづけの選出方法(LC-MS/MS) $1/X^2$ vs $1/X$

①濃度(x)と測定値(y)の標準偏差(σ)が比例関係の場合、誤差の分布は $\varepsilon \sim N(0, (x_i\sigma)^2)$

= 誤差 ε の分散は $x_i^2\sigma^2$

⇒各濃度の分散を等分散にするには

濃度の二乗 x_i^2 で割ればよい。

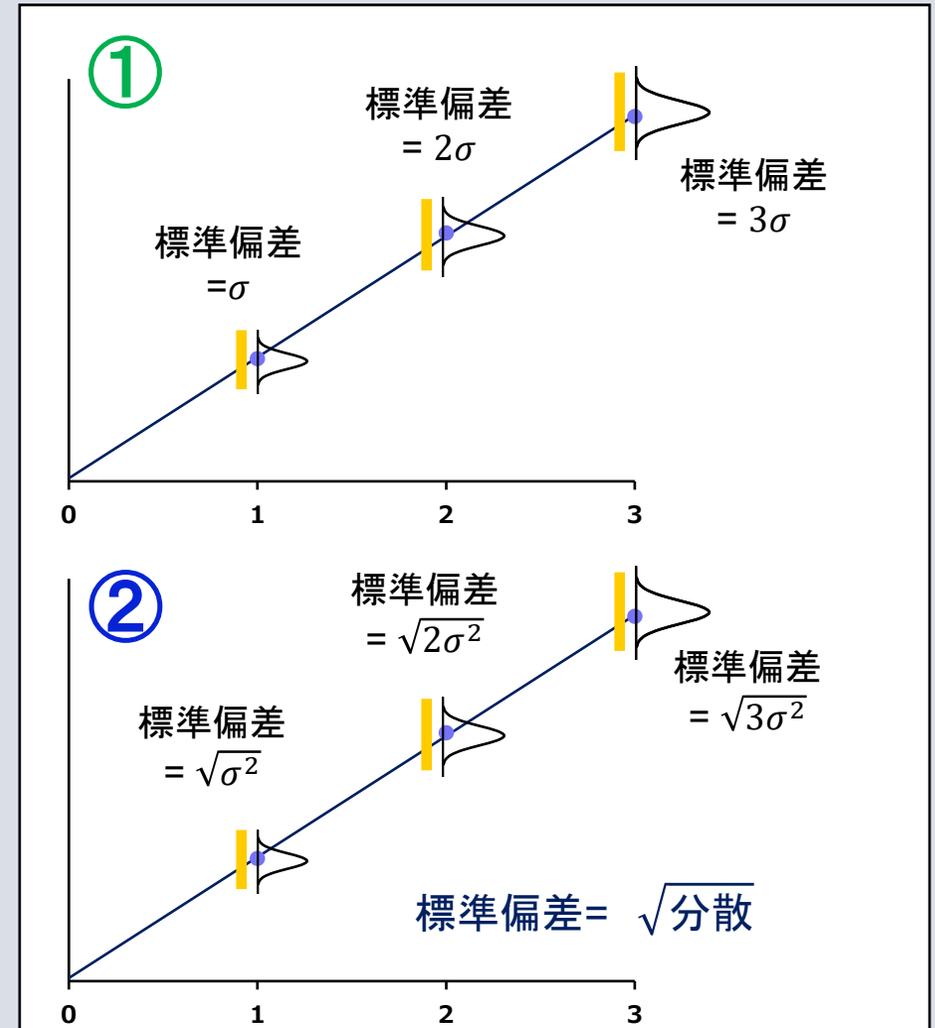
②濃度(x)と測定値(y)の分散(σ^2)が比例関係の場合、誤差の分布は $\varepsilon \sim N(0, x_i\sigma^2)$

= 誤差 ε の分散は $x_i\sigma^2$

⇒各濃度の分散を等分散にするには

濃度 x_i で割ればよい。

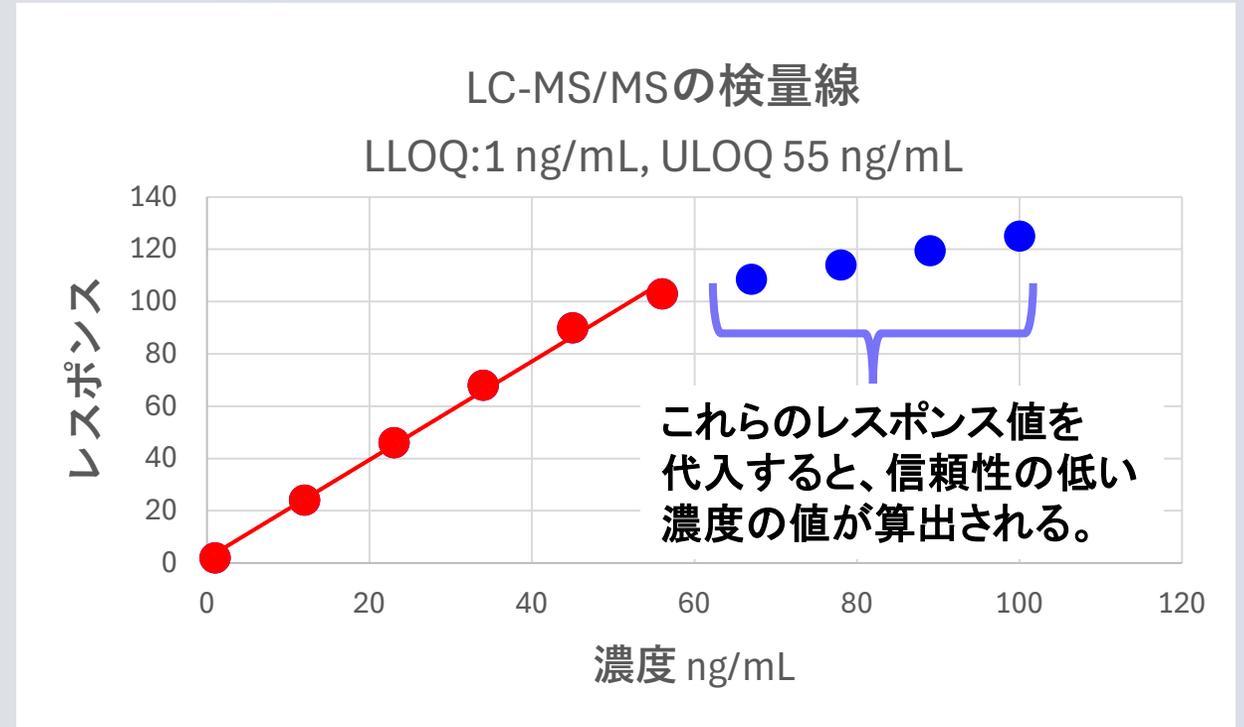
実務上は当てはまりの良い重みづけを選択することが一般的



LLOQとULOQの統計学的な意義

回帰直線を作成した範囲の外側の値を代入する操作(統計用語で外挿と呼ぶ)をすると、算出値の信頼性は低い可能性がある。

分析の定量下限LLOQと定量上限ULOQは、視点を変えると検量線に値を代入できる区間の限界値、すなわち外挿を認めないポイントとも考えられる。





クロスバリデーションで 使用する統計解析

クロスバリデーション ~ICH M10で示された目的~

クロスバリデーション(Cross Validation)とは、異なる分析法や施設間での分析結果の関係性を確認するために行われる手法のこと。

ICH M10で示されたクロスバリデーションの目的は、

複数のPK分析法及び／又は複数の分析施設が関与している場合に報告されたデータがどのように相関しているかを示すことである。

6.2 クロスバリデーション

クロスバリデーションは、複数の生体試料中薬物濃度分析法及び／又は複数の分析施設が関与している場合に、報告されたデータがどのように相関しているかを示すために必要である。

中略

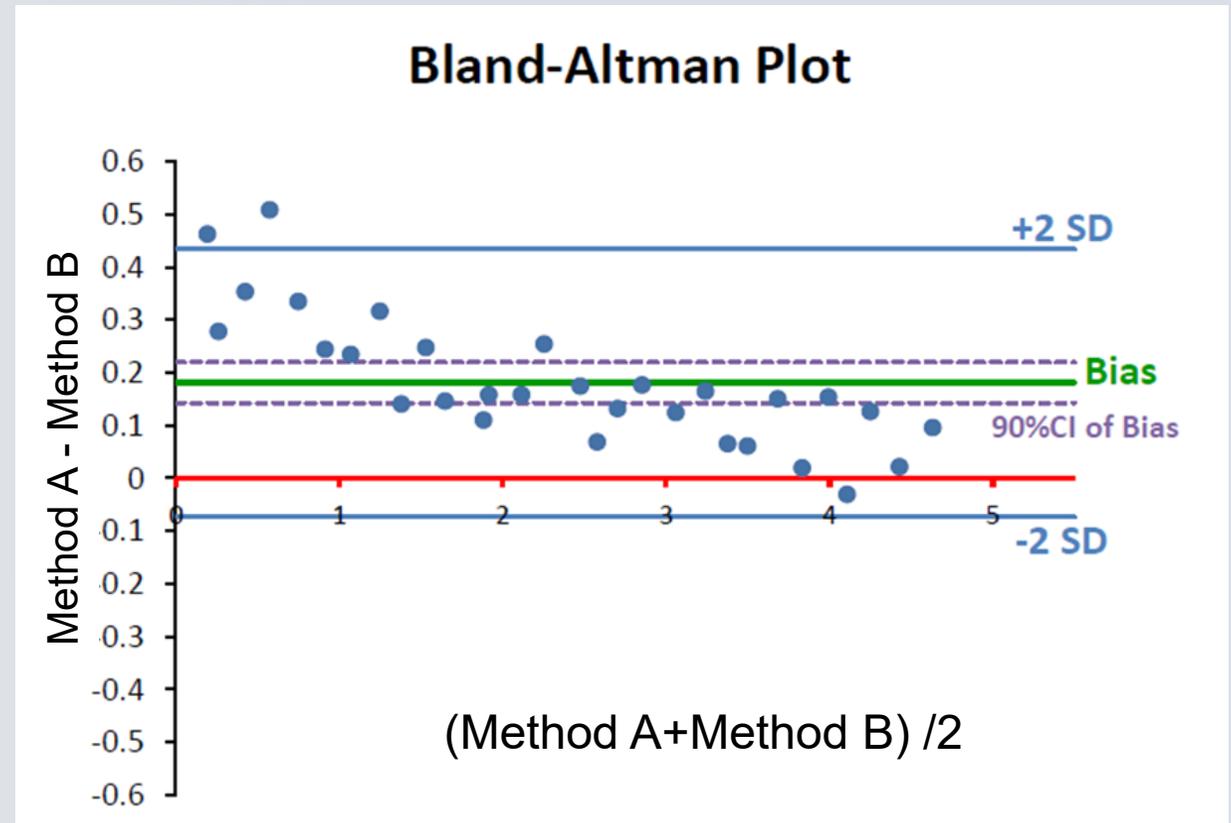
バイアスは、Bland-Altman プロット又は Deming 回帰により評価することができる。その他に二つの分析法間の一致度を適切に評価できる方法（例えば、一致相関係数）を使用してもよい。あるいは、バイアスを評価するその他の方法として、各分析法を用いて測定した実試料の濃度を時間に対してプロットすることも可能であろう。

ICH M10 6.2.クロスバリデーションより抜粋

分析法の比較①: Bland-Altmanプロット

同じ検体の分析法Aと分析法Bの結果の平均をx, 結果の差(バイアス)をyとしてプロットする手法である。

- バイアスの値や傾向を
詳細に確認できる
- 補助線として**信頼区間**などを
示すことが可能



信頼区間: ある統計量が真の値を含む範囲のこと

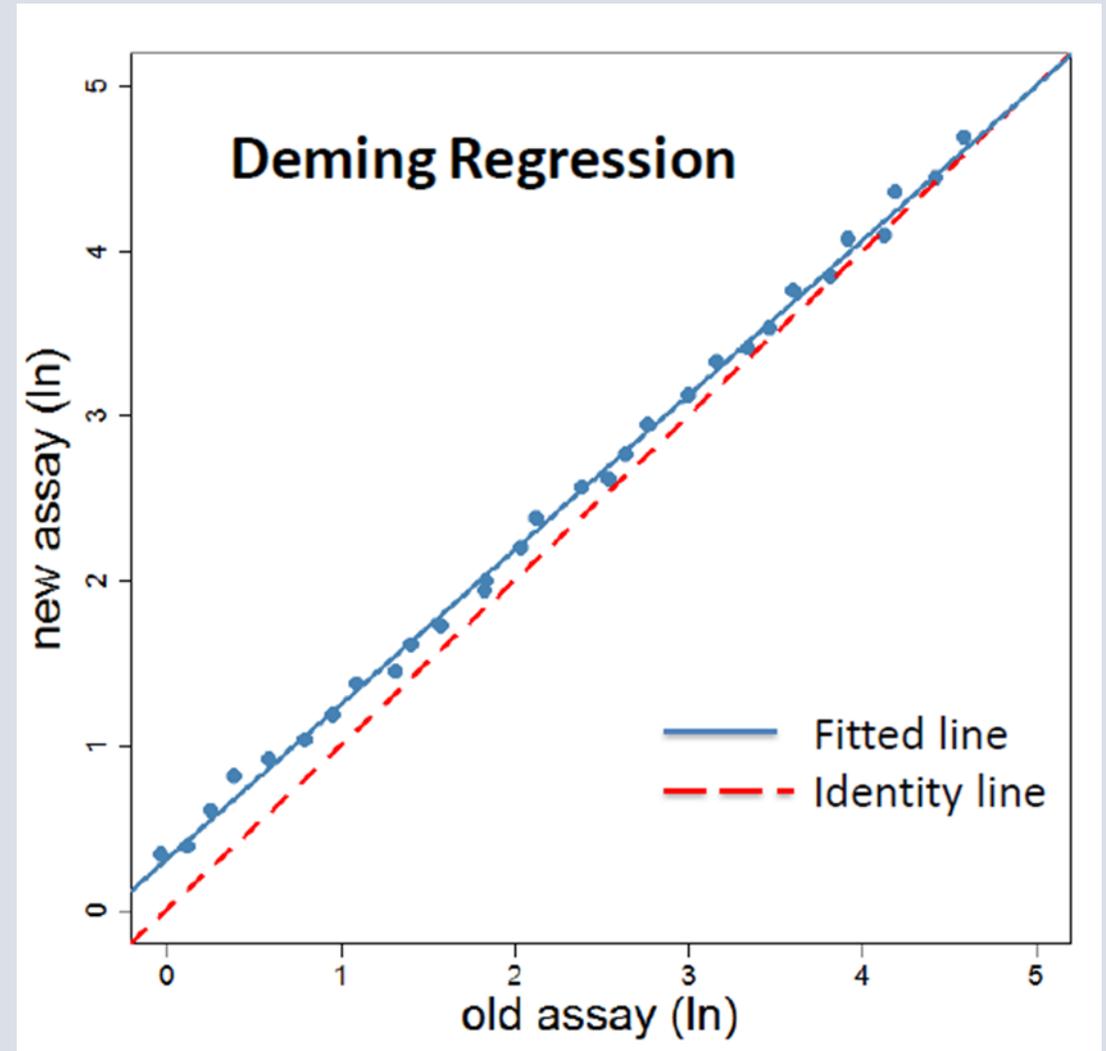
例えば95%信頼区間は、100回の観察のうち95回はその区間内に真の値が含まれることを意味する

分析法の比較②: Deming回帰

同じ検体の分析法Aの結果をx、分析法Bの結果をyとしてプロットし、線形モデルを作成する回帰分析の一種である。

この際、得られた線形モデルを2種の分析法の分散の比で補正する。

- 線形モデルとデータの全体的な関係性を可視化できる
- 線形モデルが外れ値に影響されづらい



Bland-AltmanプロットとDeming回帰のPros/Cons

解析法	Pros(長所)	Cons(短所)
Bland-Altmanプロット	<ul style="list-style-type: none">データペアの相対的な関係性やバラツキの傾向が理解しやすい信頼区間を平均に対して並行に引くことができる ⇒データが信頼区間の範囲内か否かが分かりやすい	<ul style="list-style-type: none">回帰直線は得られず、両分析の相関の程度が把握できない(平均とバイアスのプロットであるため)各々の分析法の値をグラフから読み取れない
Deming回帰	<ul style="list-style-type: none">R^2値や回帰直線が得られる。相関の程度を把握しやすい回帰直線に対する各値の離れ具合を理解しやすい回帰直線が外れ値に影響されづらい	<ul style="list-style-type: none">信頼区間を可視化しづらい ⇒濃度x、レスポンスyが大きくなる程、信頼区間が広がる補正方法が少々複雑



JBFサポーターを対象としたアンケートの結果

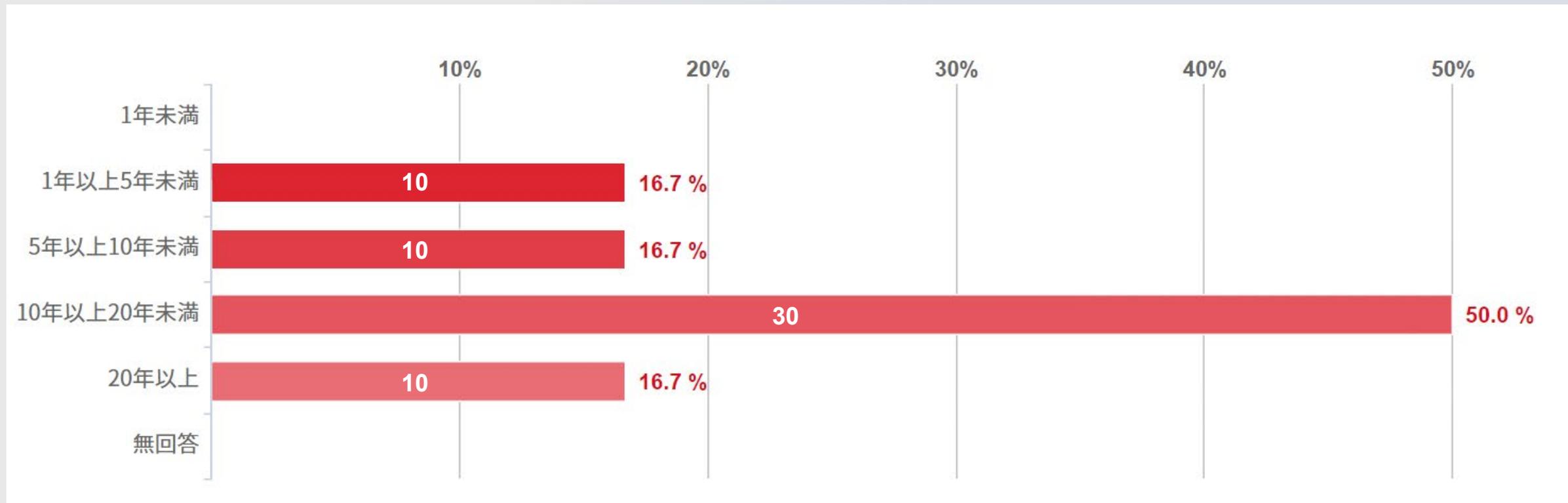
JBFサポーターを対象としたアンケートについて

本DGでは、バイオアナリシスをご担当されているJBFサポーターの皆さんが日々の業務でどのような統計関連の疑問・不明点・悩みを抱いているか確認すべく、アンケートを実施しました。

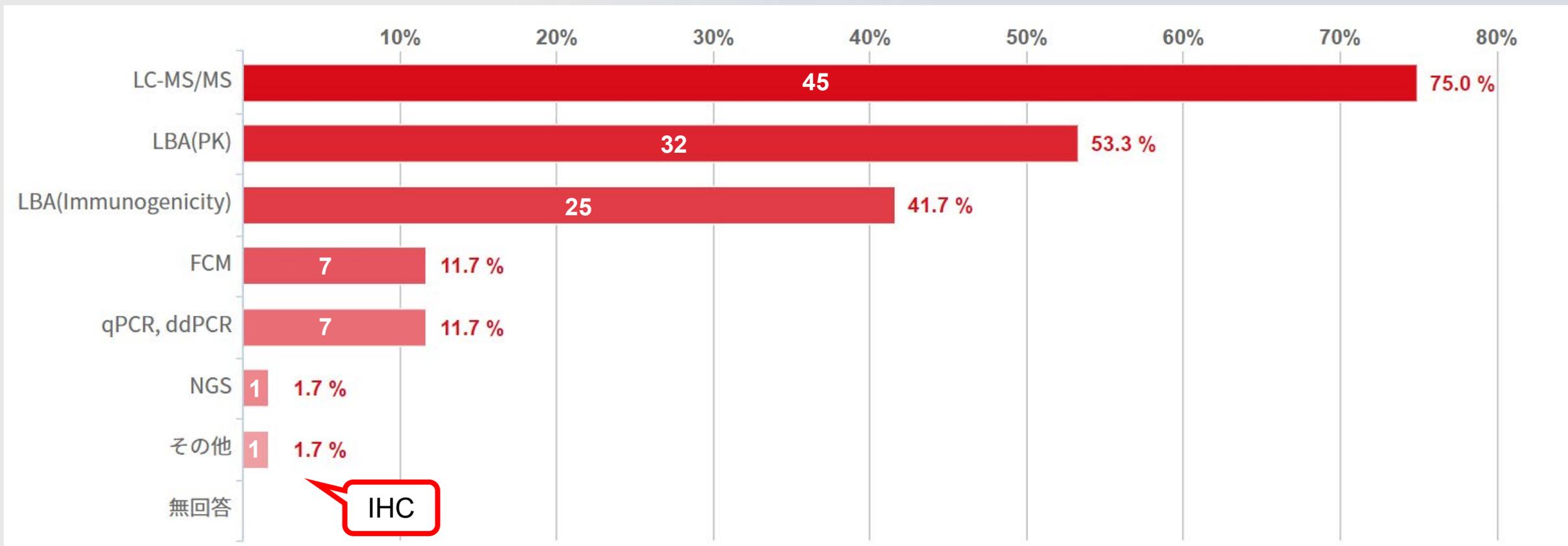
対象者: JBFサポーター(回答者の内訳はQ1とQ2参照)
集計期間: 11月12日(火)の9:00~11月26日(火)の18:00
設問数: 7問
回答者: 60名(重複回答なし、(回答者の内訳はQ1とQ2参照))

ご回答いただいた方々に、この場をお借りして御礼を申し上げます。

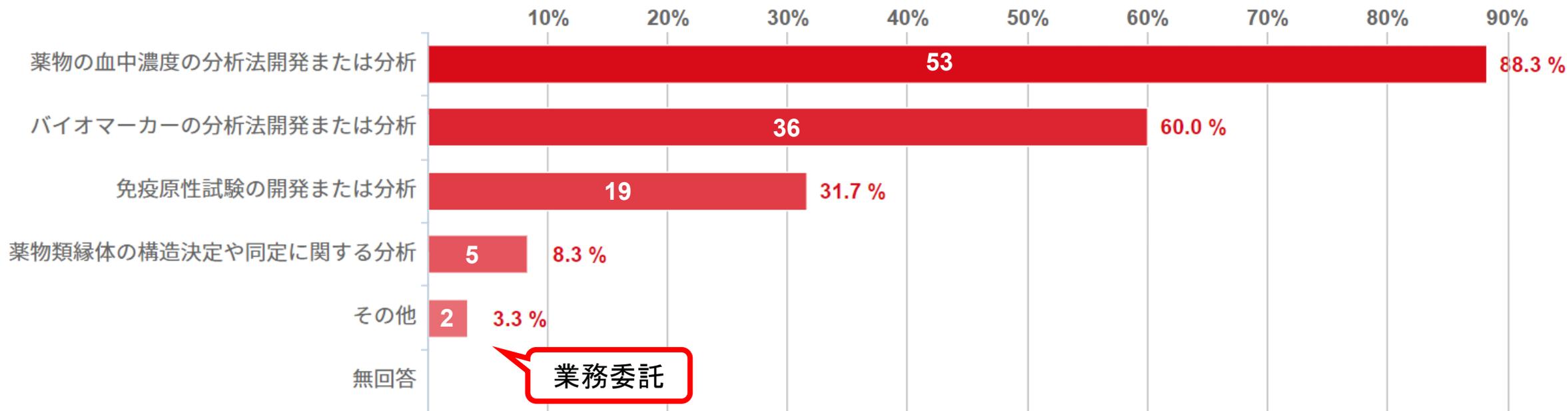
Q1. あなたのバイオアナリシス業務の経験年数を選択してください。



Q2. あなたが普段取り扱う、 分析プラットフォームを選択してください。複数回答可能です。



Q3. あなたが現在、実施しているバイオアナリシス業務 を選択してください。複数回答可能です。



Q4. バイオアナリシスに必要な統計用語や統計関連の知識として、下記以外で何があると考えられますか。

事前に示した統計用語

有効数字、平均、偏差、平均値(算術平均、幾何平均、調和平均)、中央値、最頻値、分散、標準偏差、変動係数、バイアス、相対誤差、真度、精度、散布図、ヒストグラム、箱ひげ図(四分位数、四分位範囲)、ブランド-アルトマンプロット(Bland-Altman plot)、単回帰、デミング回帰(Deming regression)、パッシング-バブロク回帰(Passing-Bablok regression)、決定係数、相関係数、t分布、 χ^2 乗分布、分散分析(ANOVA)、正規分布、標準正規分布、母平均、標本平均、母分散、不偏分散、自由度、正規性、信頼区間、有意水準、両側検定、片側検定、第1種の誤り、第2種の誤り、p値、歪度、尖度、共分散、スピアマン係数、スミルノフ=グラブス検定(Smirnov-Grubbs' test)、ディクソンの棄却検定(Dixon's outlier test)、ルビーン検定(Levene test)、クラスカル=ワリス検定(Kruskal-Wallis test)、ブラウン-フォーサイス検定(Brown-Forsythe test)、シャピロ-ウィルク検定(Shapiro-Wilk test)、ポアソン分布、各種グラフの使い分けと読み方、フィッシャーの3原則の理解、サンプルサイズの適切な計算方法、正規分布でない分布を正規化する方法、外れ値とその扱い、ADA分析のCut point算出法の理解、シグモイドカーブのフィッティングの方法(例えば、信頼区間の算出の仕方、パラメータの意味、Emaxモデル等)、薬物動態の理解、同等性(BA/BE)の評価、代謝物の標品がない状態でどうするか?等の判断、検量線の重み付けの理解

Q4. バイオアナリシスに必要な統計用語や統計関連の知識として、下記以外で何があると考えられますか。(続き)

ご提案いただいた統計用語・知識

t-SNE, UMAP(次元削減・次元圧縮), FlowSOM(クラスタリング解析)

群間の有意差検定に用いる統計手法の選択と決定樹の設定

乖離度(ISRの際に算出する)

検量線の回帰方法の選択(Linear、Quadratic、Powerなど)

最小二乗法

主成分分析, 多変量解析(PLS回帰など)

偽陽性率の両側XX%(例えば99%)信頼区間

変数変換(対数変換など)

4 parameter logistic curveの理解

ROC曲線・感度・特異度(バイオマーカーの場合)

同等性や線形性の許容範囲及び判断基準

薬物動態パラメータ

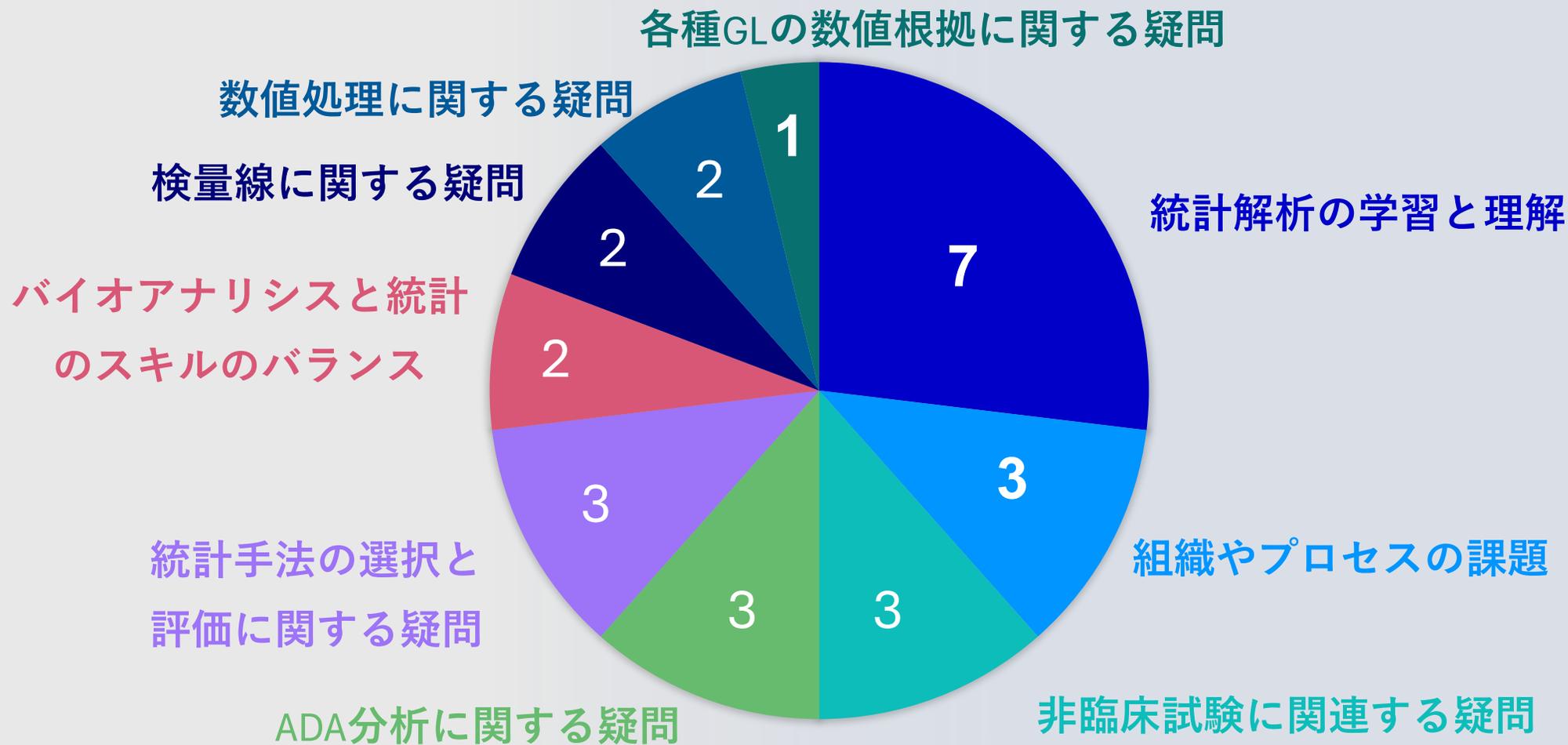
その他のコメント

- 検定については各種用語や内容を理解も大切だが、どの検定を用いるべきなのか判断できることの方が重要。
- 大方網羅されていると思うが、それで正しいか自信が持てないのが現状であり課題。

Q5. 普段のバイオアナリシス業務において、統計関連の悩みやご不明点が発生した場合、それをどのように解決しますか。複数回答可能です。



Q6. バイオアナリシス業務で感じている、統計解析に関する悩みやご不明点があれば、自由に記述してください。



Q6. バイオアナリシス業務で感じている、統計解析に関する悩みやご不明点があれば、自由に記述してください。

本ポスターのP88~94が
参考になれば幸いです

統計解析の学習と理解 7件

- 統計解析は何を、どこまで、どのように学習すればいいかわからず、自己学習のハードルが高い。
- 統計解析の理解に乏しいため、文献の通りに計算しても、なぜそうするのか理解できないことがある。
- 解析に使用している統計解析が必要十分かどうかの判断に困る。そもそも統計知識があまり無いと感じながら実務を行っている。
- 統計解析を理解できておらず、委託先の知識・経験に頼っている。統計解析というものへのハードルが高い。理想的には統計解析を理解して業務に当たるべきだと思っているが、できていないのが実状。統計解析を身近なものに感じて業務に取り組みたい。
- どこまで理解していなければならないのかが体系的に整理できていない。学習するうえでの参考書や資料等を知らない。
- 分析に必要な統計の知識がどのような範囲なのかがよくわからない。特にガイドラインに関連することであれば、何らかの分析の統計解析に関する指針も必要ではないか。
- 統計手法や統計処理で得られた値について、本質的な意味を理解して使用、選択したいが、論文等で調査しても理解することが難しく、理解できないまま使用してしまっていることが多い。

Q6. バイオアナリシス業務で感じている、統計解析に関する悩みやご不明点があれば、自由に記述してください。

バイオアナリシスと統計のスキルのバランス 2件

- バイオアナリシスと統計をバランスよく理解している人が少ない。重要な判断をするデータとなることもあり、社外に気軽に聞きにくいけれど、詳細を話さないと正しく相談できないというジレンマがある。
- バイオアナリシスに即した統計学を学べる・相談できる相手が社内には存在しない。

組織やプロセスの課題 3件

- 教えてもらってもピンとこない。良し悪しを判断出来ない。
- 自社で言えば、分析も出来て統計解析も出来てという人材が少ないこと。
- 数値をこねくり回して恣意的により有利な結果を導き出すというようなことが無いように、透明性を持ち、かつ誰にでもわかりやすいプロセスを構築するのは非常に難しいと思う。

同様の悩みを抱えている方も多くいらっしゃるようです。本ポスターのP88～94が参考になれば幸いです

Q6. バイオアナリシス業務で感じている、統計解析に関する悩みやご不明点があれば、自由に記述してください。

ADA分析に関する疑問 3件

- ADAのカットポイントの算出の手順。解析するごとに毎回悩んでしまう。
- ADA分析のCut point算出法の理解が正しくできているか不安になる。
- (ADA分析のカットポイント設定において)複数の集団のバラツキや平均を比較したいとき。

1) 11th JBF Symposium DGポスター <https://bioanalysisforum.jp/common/pdf/event/dg/DG2019-43.pdf>

2) 15th JBF Symposium 基礎講座 https://bioanalysisforum.jp/images/2024_15thJBFS/D1-A4-02_JBF15_Tamiki_Mori.pdf

ADA分析のCP設定の簡単な解説は本ポスターP25~30や下記のJBF資料、CP設定に関する議論は本ポスターP74~76などが参考になれば幸いです

検量線に関する疑問 2件

- 検量線の重み付けについて正しく理解したい。
- 適切な重みづけの選択を行うための統計学的な解説(慣習的に1/xを選択することが多いが、統計学的な根拠があれば知りたい)。

本ポスターP38,39,71を是非ご参照ください

Q6. バイオアナリシス業務で感じている、統計解析に関する悩みやご不明点があれば、自由に記述してください。

非臨床試験に関連する疑問 3件

- 非臨床部門に所属しているが、安全性マージンを求める際に適切な臨床AUC(あるいはCmax)は幾何平均、算術平均、中央値のどれが妥当か？

Tmaxを除くPKパラメータでは対数正規分布をとることが多く、幾何平均を算出するのが一般的。
詳しくは[生物学的同等性ガイドライン](#)を参照

- 非臨床TK測定等におけるn数の設定方法。動態関連の試験では最低限SDを算出できるn=3に設定することが多いが、統計的には濃度値のバラつき等を考慮して決定すべきなのか。

個体内・個体間の濃度のバラつきを考慮すべきだが、[TKに関するガイダンス](#)には、「通常、TK試験では統計学的な意味での高い精度は必要とされない」とあるのでSDを算出可能なn=3に設定するのが一般的と考える

- TK測定における群間の濃度値の増減を考察する際に、ガイドラインに記載の2倍の差を基準に判断しているが、統計的に適切ではない気がしている。

各種ガイドラインの数値根拠に関する疑問 1件

本ポスターP68~70, 81をご覧ください

- 真度、精度の考え方や判断基準のガイドライン上の基準ではなく、統計学的な考え方。

Q6. バイオアナリシス業務で感じている、統計解析に関する悩みやご不明点があれば、自由に記述してください。

社内外関係者とよく協議しながら、目的に合致した解析手法を選択する必要があります。クロスバリデーションに関しては本ポスターP77~81をご覧ください

統計手法の選択と評価に関する疑問 3件

- クロスバリデーションの評価方法(ひとまずBland-Altmanプロットを実施するが、偏りの判断目安をどうするか、偏りがあったときにどう対応するか)。
- 自身が選択した解析方法が明確に正しいかが分かりにくい。確証が持てない。
- 用いている解析手法が適しているのかが不明。その手法で結果が得られなかった場合の代替案が不明。

有効数字に関しては本ポスターP67をご参照ください

数値処理に関する疑問 2件

- 数字を丸める(有効数字にする)際に、どのタイミングでどのように丸めるのが正しい(恣意的な処理とはならない)か迷うことがある。
- 正規化を目的として対数変換して解析するべきか否か、迷うときがある。

Q7. 最後に我々のDG[バイオアナリシスに関連する統計学]に議論してほしいこと・成果物(基礎講座やポスター発表など)に説明をしてほしいこと等があれば、自由に記述してください。

数値処理に関する疑問

統計解析で使用する
ソフトに関する疑問

統計手法の選択と
評価に関する疑問

各種ガイドラインの数値根拠に関する疑問



統計解析の学習と理解

ADA分析に関する疑問

Q7. 最後に我々のDG[バイオアナリシスに関連する統計学]に議論してほしいこと・成果物(基礎講座やポスター発表など)に説明をしてほしいこと等があれば、自由に記述してください。

統計解析の学習と理解 6件

- 統計解析でやってはいけないことを挙げていただけると参考になると思う。
- 統計検定を受験する意味はあるか。何級までであると良いか？
- 統計値を算出する数式一覧。統計学の教科書の場合、とても長い説明や式の変換も書いてあり、実用面ではわかりにくく感じることもある。
- 何から手をつけてよいかわからないときに、スタンダードな方法、基にしている論文等の情報を知りたい。
- 日常的に使用しているバイオアナリシスの統計を正しく理解できる説明などが欲しい。
- バイオアナリシスと統計のスキルのバランスを知りたい。

全てに回答出来ず恐縮ですが、本ポスターP81~94に参考になる書籍や本ポスターで活用した論文を用意しました

数値処理に関する疑問 2件

- 数字を丸める(有効数字にする)際に、どのタイミングでどのように丸めるのが正しいか迷うことがある。
- 外れ値の判定、取り扱い方を理解したい。

有効数字に関してはP67、外れ値に関してはP28,72,73をご覧ください

Q7. 最後に我々のDG[バイオアナリシスに関連する統計学]に議論してほしいこと・成果物(基礎講座やポスター発表など)に説明をしてほしいこと等があれば、自由に記述してください。

ADA分析のCP設定の簡単な解説は本ポスターP25~30や下記のJBF資料、CP設定に関する議論は本ポスターP74~76などが参考になれば幸いです

ADA分析に関する疑問 5件

- ADAのカットポイントの算出に統計が用いられるが、どれがゴールドスタンダードになるのか知りたい。
- ADAカットポイントの解析やクロスバリデーションの解析等、統計が関わってくる規制文書の内容に対して、どのような対応をどのような根拠を元にどのような考えで行ったか、といった内容。
- ADA分析のCut point算出法をデータ例を使って動画にして公開すること。
- ADAのカットポイントの設定に関する説明をしていただきたい。
- 外れ値の評価や陽性対象の選択など、測定法検討時から実試料測定まで複数の統計手法が関わっている。その辺りを具体例を基に説明する基礎講座があれば、とてもありがたい。

1) 11th JBF Symposium DGポスター <https://bioanalysisforum.jp/common/pdf/event/dg/DG2019-43.pdf>

2) 15th JBF Symposium 基礎講座 https://bioanalysisforum.jp/images/2024_15thJBFS/D1-A4-02_JBF15_Tamiki_Mori.pdf

Q7. 最後に我々のDG[バイオアナリシスに関連する統計学]に議論してほしいこと・成果物(基礎講座やポスター発表など)に説明をしてほしいこと等があれば、自由に記述してください。

重みづけに関してはP38,39,71、クロスバリデーションの解析に関してはP41~45, 77~81をご覧ください

統計手法の選択と評価に関する疑問 3件

- 適切な重みづけの選択を行うための統計学的な解説が欲しい(慣習的に $1/x$ を選択することが多いが、統計学的な根拠があれば知りたい)。
- クロスバリデーションの評価方法について、ひとまずブランド-アルトマンプロットを実施するが、偏りの判断目安をどうするか、偏りがあったときにどう対応するか。
- 最近の話題でいうとM10のCrossValiの指標として示されたBlandAltmanとDemingRegressionのpro/Conを知りたい。

統計解析で使用するソフトに関する疑問 3件

- 使用している解析ソフトの種類など。
- 最近のexcelでは信用できるものと考えて良いのかどうか。
- 統計学を楽に使いこなせるツールの紹介(pythonやRプログラム含め)。

ExcelやR、分析装置の解析ソフトで十分に簡易的な解析は可能と考えます。
※Shapiro-Wilk検定やPassing Bablok回帰などの複雑な処理や審査報告書に記載する解析はR, SASなどが必要な場合があります

Q7. 最後に我々のDG[バイオアナリシスに関連する統計学]に議論してほしいこと・成果物(基礎講座やポスター発表など)に説明をしてほしいこと等があれば、自由に記述してください。

各種ガイドラインの数値根拠に関する疑問 4件

- ICHM10のガイドラインなどで具体的なクライテリアの記載がない部分について、統計的に適切なクライテリアを提示できそうな部分があったら、提案してほしい。
- QCの15%とか20%基準とかの数値にどういう意味があるのか (なぜLCMSは20%ではだめなのか)。
- 特に統計的な手法が必要と思われる解析や基準設定などについて、事例を紹介してはどうか？具体的にその有用性などを実感できる、あるいは必要であることが目視できるような事例の提示があるとわかりやすい。
- 真度、精度の統計学的または薬物動態学的な判断基準について。

**各ガイドラインに記載の真度・精度のAcceptance Criteriaの基準値やISRの実施方法に統計的根拠があるのかを議論しました。
本ポスターのP68～70, 81がご参考になれば幸いです**



DGでの議論内容の ご紹介

DGでの議論に関する概要について

本DGでは、統計解析を専門とするオブザーバーを交えながら、

- ① メンバーがバイオアナリシス業務で感じている統計関連の疑問・不明点・悩み
- ② JBFサポーターを対象としたアンケートで集計された、サポーターの皆様が日々のバイオアナリシス業務で感じている統計関連の疑問・不明点・悩み

を題材とした議論を実施しました。

多数の議論を会議中に実施しましたが、一部を抜粋して紹介します。

これらはメンバー個人およびDGの意見・考えであり、

JBFとしての意見ではないことをご理解ください。

Q1. 有効数字を3桁とするということは、 「4桁以降の数字に意味がない」ということを示しているか？

社内の臨床薬理に携わる者から「有効数字3桁とするということは4桁以降の数字に意味がない、ということか」という質問が挙がった。どのように答えるのが適切か？

4桁目に意味がないというわけではなく、3桁目までが信頼できる情報であることを示している。

なお、分析結果の有効数字はプロセス全体の操作精度に依存する。例えば、分析装置の感度が高くても、使用した天秤やメス容器の精度が低ければ、データの精度は低い方に依存する。

また、有効数字3桁は0.1%の精度を示し、4桁は0.01%の精度を示すが、そのような微小な差で臨床的な差異が現れることは極めて稀である。

4桁の精度が必要な場合、分析プロセス全体を精密にする必要があるが、バイオアナリシスではそのような状況はほとんどないものと推察される。

有効数字3桁の場合

0.1 2 3 4 5 6...

信頼できる範囲

測定精度の限界を
超過しており信頼できない

Q2. ICH M10の真度・精度のAcceptance CriteriaやQCの分析数、ISRのPass/Failの基準などに統計的根拠はあるのか？

ICH M10¹⁾には、PK分析のバリデーションや実試料分析における真度・精度のAcceptance CriteriaやQCの分析数、ISRのPass/Failの基準などの推奨値が示されている。これらの値は、例えば、区間推定の考えや統計的有意性といった統計的な根拠を基礎として設定された数値なのか？

ICH M10の真度・精度のAcceptance CriteriaやQCの分析数、ISRの基準等は、1990年から始まったCristal City会議^{2)~4)}やAAPSのワーキンググループでの議論^{5), 6)}を基に、複数機関のコンセンサスを得て決定された。

ただし、Cristal City会議のWhite Paper等^{2~6)}からは、これらの基準値や方針に統計的根拠があったかどうかは確認できなかった。

そのため、質問の回答としては「不明」となるが、これらの基準値や方針は実験的事実や技術的に想定される一般的な状況を考慮して議論されている。

また、各専門家から合意を得られたものである以上、基準値・方針としては十分妥当なものと考えられる。

一方、2018年のEBFのWhite Paper⁷⁾では、こうした実験的事実の視点だけではなく、技術に依存しない一貫した基準を確立し、Acceptance Criteriaを再考すべきと指摘している。

- 1) ICH. (2022). M10 Guideline.
- 2) Shah, V. P. et al. Pharm. Res. 1992, 9, 588-592.
- 3) Viswanathan, C. T. et al. AAPS J. 2007, 9.
- 4) Fast, D. M. et al. AAPS J. 2009, 11, 238-241.
- 5) Miller, K. J. et al. Pharm. Res. 2001, 18, 1373-1383.
- 6) DeSilva, B. et al. Pharm. Res. 2003, 20, 1885-1900.
- 7) Timmerman, P. et al. Bioanalysis 2018, 10 (16), 1255-1259.

Q3. 分析におけるn数の設定や有効数字の設定の仕方はどうすればよいのか？

実試料分析を実施するにあたり、n数の設定などを変更する必要があった場合どうすればよいか、また有効数字の桁を決定するのも悩むことがある。

PK分析のバリデーションや実試料分析では、基本的にICH M10¹⁾に基づいてn数やAcceptance Criteria、QC濃度や本数を決定することが推奨される。ただしICH M10にも記載されているように、科学的根拠があれば代替手法も認められる。

PK分析以外の分析(免疫原性分析やバイオマーカー分析など)では、状況に応じて設定方法が異なるため、一概には述べることはできない。ガイダンスがない場合や各種レギュラトリーと異なる条件を採用する場合は、統計的な論拠だけでなく、臨床薬理の観点からデータの解釈が可能かどうかを社内外の関係者で十分に考察することが重要である。

有効数字も同様に、臨床薬理的な考察が可能な精度の分析を追求し、関係者とよく相談して設定することが推奨される。

生体試料中薬物濃度分析法のバリデーションの目的は、その分析法が使用目的に適していることを立証することにある。本ガイドラインの推奨事項からの変更は、適切な科学的根拠が示される場合には受け入れられる可能性がある。申請者が代替の方法を提示又は採用する場合には、分析法バリデーションの方法に関する重要な変更について規制当局と相談することが望ましい。

ICH M10 Step5¹⁾
1.1 目的 より引用

1) MHLW(2025), ICH M10 生体試料中薬物濃度分析法バリデーション及び実試料分析 (Step 5)

Q4. ISRやランのAcceptance Criteriaで頻出する「すべてのサンプルの2/3が適合する」という基準の根拠は何に由来するか？

ICH M10では、分析単位の判定基準において、「すべての QC 試料の少なくとも3分の2、かつ各濃度における少なくとも50%が理論値の±XX%以内でなければならない」との記載がある。またISRの項においても、「試料数の少なくとも3分の2において、乖離度が±XX%以内でなければならない」との記載がある。この「3分の2以上が適合する必要がある」という基準は何に由来するものなのか？少々緩すぎないだろうか？

分析単位の判定基準は1990年開催のCristal City I会議¹⁾、ISRの基準は2006年開催のCristal City III会議²⁾(具体的な基準は1999年のFDAガイダンスの分析単位の判定基準を参考にしたものと推察される)で複数の専門家の合意の下、推奨されたものである。

分析性能の基準に関しては、

- アッセイのバイアスや変動性が大きい場合は誤ってFailと判断する率が高くなり、分析の性能が劣化していなくても、分析をFailさせてしまう可能性があること

ISRの基準に関しては、

- 分析数がISRの結果に大きな影響を与える可能性があり、(推奨されたサンプル数であっても)分析数が少ないとISRの失敗確率が高まること

などが論文^{3), 4)}により指摘されている。

1) Shah, V. P. et al. Pharm. Res. 1992, 9, 588-592. 2) Fast, D. M. et al. AAPS J. 2009, 11, 238-241.

3) Kringle, R. O. Pharm. Res. 1994, 11, 556-560. 4) Hoffman, D. AAPS J. 2009, 11 (3), 570.

Q5. 回帰直線の重み付けをする際、濃度xとレスポンスyのどちらを重みづけするか？

回帰直線の重みづけにおいては、濃度xとレスポンスyどちらかを重みづけをするべきか、 $1/x^2$ (or y^2), $1/x$ (or y), 重み付けなしのどれを選択するべきか、という複数の組み合わせが存在すると考える。これらをすべて組み合わせで試して、最も決定係数が1に近い(残差平方和が小さい)ペアを選択しているのだが、その手法で間違いはないのだろうか？

LC-MS/MSの検量線は $1/x^2$ (濃度の2乗の逆数)、LBAの検量線は $1/y^2$ (レスポンスyの2乗の逆数)で重みづけするのが一般的^{1),2),3)}。

この差異は、LC-MS/MSとLBAの検量線がそれぞれ線形、非線形であるためである。

想定される組み合わせをすべて試して、最も決定係数が1に近い(残差平方和が小さい)ペアを選択するという方法も間違いではない。

1) Gu, H. et al. Anal. Chem. 2014, 86 (18).

2) Azadeh, M. et al. AAPS J. 2018, 20 (1), 22.

3) Xiang, Y. et al. AAPS J. 2018, 20 (1), 45.

LC-MS/MS(線形)

低濃度域では信号が弱く、ノイズの影響でばらつきが大きくなる。高濃度域では信号が飽和しやすく、ばらつきが大きくなる。濃度xを用いた重みづけ(例: $1/x$ や $1/x^2$)で、濃度範囲全体のばらつきを均一にし、正確な回帰直線を得る。

LBA(非線形)

高濃度域では、濃度が増加してもレスポンスがそれほど増加せず、レスポンスのばらつきが大きくなる。濃度に重みを付けても影響を軽減できないため、レスポンスyを用いた重みづけ(例: $1/y$ や $1/y^2$)で、測定値のばらつきを補正し、非線形性の影響を減少させる。

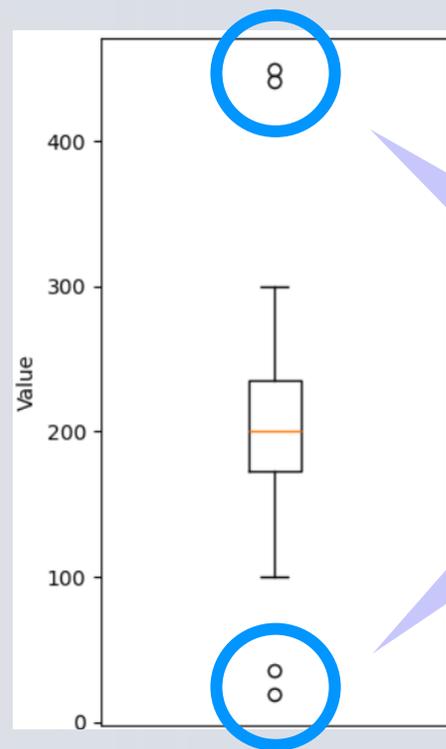
Q6. あるデータを外れ値と解釈することと、 データから除外することの違いは何か？

例えば、ADA分析のカットポイント設定において、NCから検出された極めて高いレスポンスを持つデータを扱う際に外れ値と解釈することとデータを除外すること、両方に若干の混乱がある。明確な線引きはあるか？

外れ値と解釈することは、データの異常を認識し、適切な対策を講じるためのステップであり、必ずしもデータを除外することを意味しない。一方、データから除外することは、分析の精度を保つための具体的なアクションである。

その外れ値を除外することは恣意的な処理であるため、悪用すればデータを見かけ上ポジティブに変えることが可能となる。

その外れ値が生物学的誤差か、それとも偶然誤差か等を繰り返し測定によって評価することが重要である。



外れ値と解釈することと、外れ値をデータから除外することには違いがある

⇒その外れ値の原因を見極めることが重要

Q7. なぜ箱ひげ図ではひげを IQRの1.5倍のところにはひくことが多いのか？

箱ひげ図のひげの長さに何か明確な数値的根拠があるのか疑問に感じた。なぜ、IQRの1.5倍なのか。

IQRの1.5倍にひげを引くことで、外れ値を特定しやすくなるため。1.5倍という基準は、1倍だと外れ値の基準が厳しすぎ、2倍だと緩すぎるため、適度なバランスを取るために採用されている。

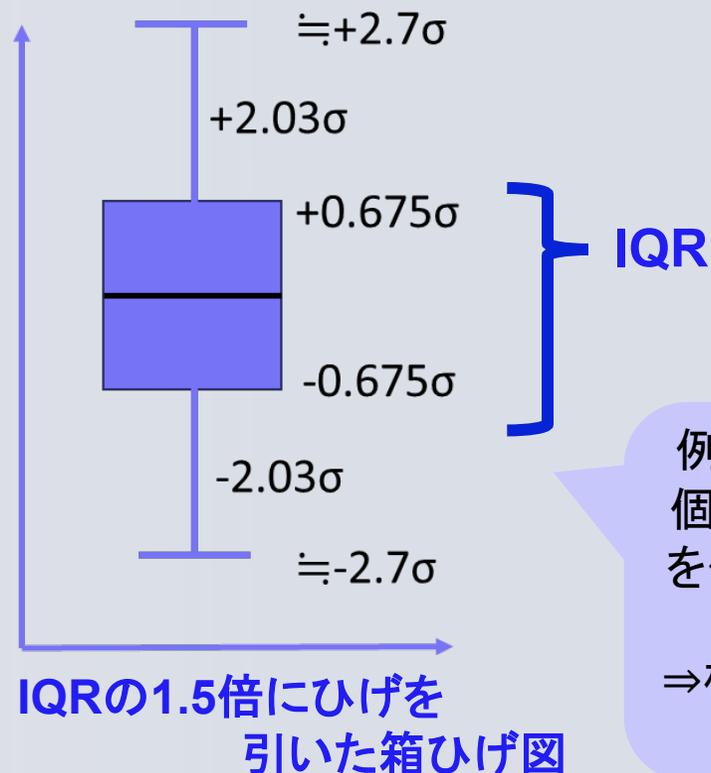
IQRは全データの50%($\pm 0.675\sigma$)に相当する

箱の幅は 1.35σ なので、 1.35σ の1.5倍にひげを引くと片方のひげの長さは 2.03σ となる

ひげの範囲は**分布が正規分布の場合**、平均値(かつ中央値)から $\pm 2.7\sigma$ の範囲に相当する

この範囲にほとんどのデータが収まるはずであり、これから外れた値は極めて稀(=外れ値)とみなせる。

ただし、正規分布の前提で定められた基準であることに注意が必要である。



例えば、正規性を持つ1000個のデータセットで箱ひげ図を作成すると、ひげの外側にあるデータは約7個
⇒極めて稀なので、外れ値とみなす

Q8. Shapiro-Wilk検定において有意水準p以上であった場合は正規性があるとは厳密に言えないのではないか？

ADA分析のカットポイント設定においては、NCで得たデータに正規性があるか否かを確認するために、Shapiro-Wilk検定を使用する。ところが本検定での結果が有意水準p以上であった場合は、帰無仮説を棄却できないということになる。これは「帰無仮説に従わないとはいえない」ということであり、正規性があるとは厳密に言えない、すなわち結論が出ないことにならないか？

指摘の通り、検定の原理に基づく、Shapiro-Wilk検定の結果が「帰無仮説に従わないとはいえない」となった場合、その変数の分布が正規性を持つことを意味するわけではない。

それでもこの検定を使用する意義は、「分布が正規性を持つ可能性がある」という評価を得ることができる点にある。

一旦、分布に正規性があると認めながら、ヒストグラムを用いた視覚的な評価やp-pプロットなどを用いた評価、その分布の歪度と尖度を用いた評価を並行して実施することを推奨する。

$$W = \frac{\left(\sum_{i=1}^n a_i x_{(i)}\right)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

帰無仮説(H0): 変数の分布は正規分布に従う

対立仮説(H1): 変数の分布は正規分布に従わない

p値が予め定めた有意水準未満であった場合、帰無仮説が棄却される
→対立仮説が採用

P値が予め定めた有意水準以上であった場合、帰無仮説が棄却できない
→「帰無仮説に従わない」といえない

Q9. 歪度と尖度は分布の正規性を評価するのに使用されるが、Shapiro-Wilk検定とどのように使い分けるのか？

歪度・尖度による正規性評価とShapiro-Wilk検定はどのように使い分ければ良いのか、それぞれのPros/Consを理解したい。

ADA分析のカットポイント設定においては、NCのレスポンスに正規性があるかどうかを評価するために、尖度・歪度を使用する方法やShapiro-Wilk検定を使用する方法が知られている。両者にはそれぞれ、以下に示すPros/Consがあると考えられる。歪度・尖度で簡単に評価した後、Shapiro-Wilk検定で裏付けする等、両方を使用することなども考えられる。

正規性評価法	正規性がある場合	Pros	Cons
Shapiro-Wilk検定	検定の式に代入した結果Wが、有意水準以上	<ul style="list-style-type: none">統計的に検定するため、客観的な判断が可能学術論文や研究で広く使用されている	<ul style="list-style-type: none">大規模なデータセットでは計算が複雑になる結果が有意水準以上の場合、帰無仮説を棄却できず、結論が得られないことがある
歪度・尖度の評価	実際のデータを基に算出した歪度が0付近かつ、(3を正規分布としたときの)尖度が3付近	<ul style="list-style-type: none">比較的簡単に計算でき、データの分布の形状を直感的な理解に役立つデータを視覚的に評価するための指標として有用	<ul style="list-style-type: none">統計的仮説検定ではないため、正規性を統計的に証明することはできない両値の解釈には主観が入りやすく、明確な基準がない場合がある

Q10. ADA分析のカットポイント設定において、t統計量を算出する際は、NCのn数を ∞ として算出するが、標準偏差の算出においては、不偏標準偏差を使用する。これはなぜ？

上述の状況では、NCのデータに正規性がある場合はまずt統計量を算出する。その際に、NCのn数を50(自由度49)ではなく、 ∞ として算出するのが一般的である。一方、標準偏差としては不偏標準偏差(n-1を分母にして平方根を取る標準偏差、標本の標準偏差を算出する際はこちらを採用)を採用することになる。n数を ∞ とすることはNCを母集団とみなしているように思えるが、不偏標準偏差を使用するのはNCを標本とみなしているため、NCに対する考え方に一貫性がないように感じる。

NCのn数を ∞ としてt統計量を算出するのは、漸近正規性を認めて値を近似しているに過ぎない。NCを母集団とみなすような操作ではないので、一貫性がないわけではない。

※n=50というサンプルサイズは十分に大きいため、

t統計量を正確に算出しても、n= ∞ の時と著しい差はない。

漸近正規性とは、サンプルサイズが非常に大きくなるときに、推定量の分布が正規分布に近づく性質のことを指す。これは、中心極限定理に基づいており、サンプルサイズが増加するにつれて、標本平均などの統計量が正規分布に従うようになることを意味する。

自由度	片側確率 α	
	0.05	0.01
49 (n=50の場合)	1.677	2.405
∞	1.645	2.326

Q11. 正規性を持たないデータセットを用いて Bland-AltmanプロットやDeming回帰を使用しても良いのか？

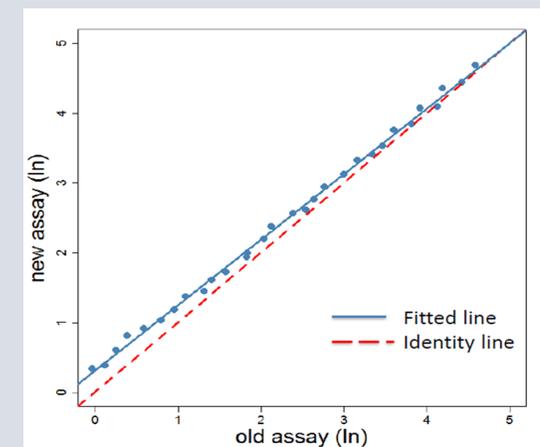
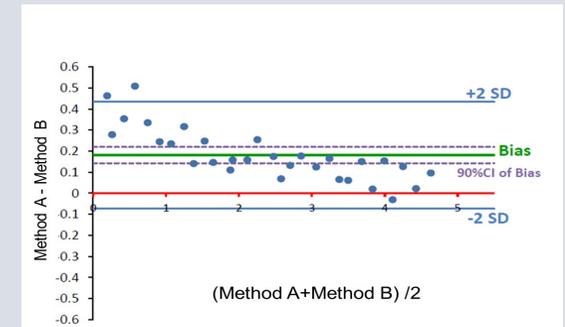
クロスバリデーションの分析法比較ではBland-AltmanプロットやDeming回帰を使用することが一般的である。通常、PK分析は調べたいポイントの薬物濃度を測定するので、正規性を持たない。一方B-AプロットやDeming回帰は正規性を持つことを前提とした解析法なので、不適切ではないか？

両手法は正規性を前提とした統計解析手法であるが、クロスバリデーションでこの手法を採用する理由の一つとしては記述統計(グラフを介してデータの性質や傾向を理解する統計手法)としての活用である。

なお、Bland-Altmanプロットにおいては信頼区間、Deming回帰においては傾きや直線の R^2 値が推測統計(データから具体的な数値を算出し、評価をする)の側面を持つ。

正規性を持つことが前提の両解析法ではあるが、正規性を持たないからといって使用してはいけないというわけではなく、正規性が低いデータセットで使用すると、回帰の説明力が低下するリスクがある、ということになる。

逆に正規性が低い状態でも、現象を十分に説明出来る(R^2 値が1に極めて近い等)のであれば、保守的に見ても現象を上手く説明できている、ということになる。



Q12. クロスバリデーションでPassing-Bablok回帰を実施するのは適切か？

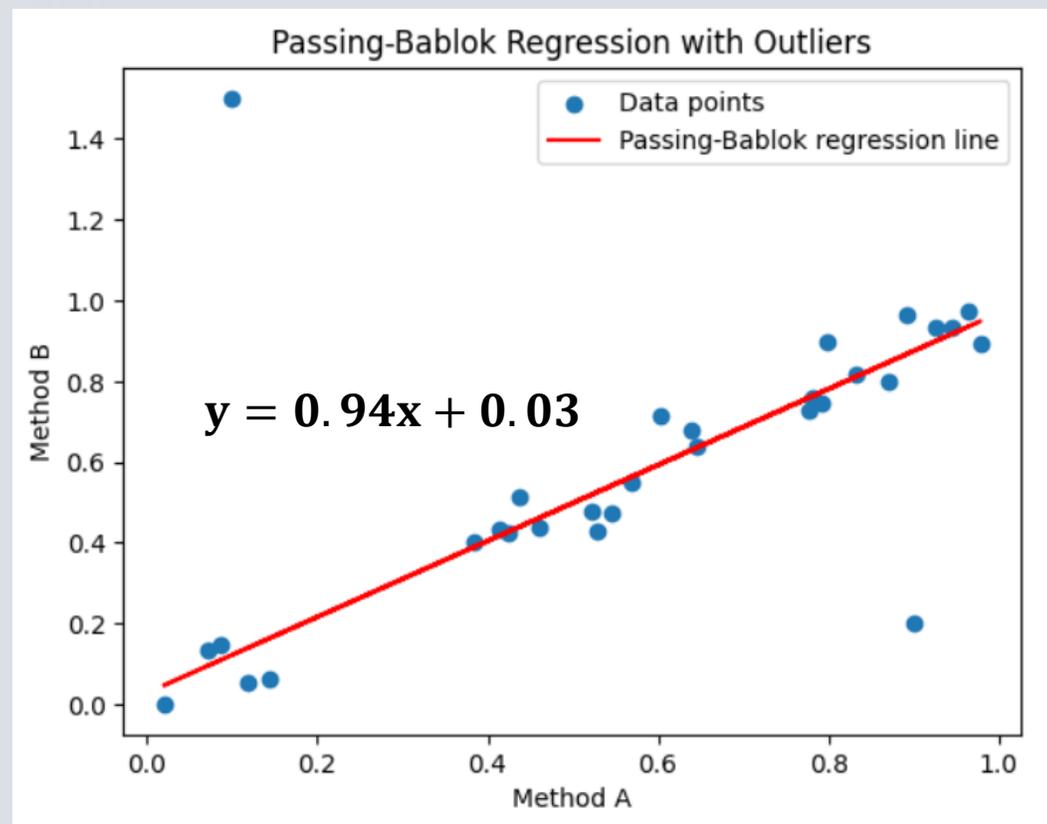
正規性を持たないデータの回帰分析には、Passing-Bablok回帰というノンパラメトリックな回帰分析が使用されることもある。バイオアナリシスのシチュエーション、例えばクロスバリデーションにおいてPassing-Bablok回帰を実施することは適切か？

この回帰は2種の分析で得たデータのペアを用いて、すべての組み合わせで傾きを求め、その中央値を採用する解析法である。仮に $n=30$ のデータのペアを得た場合、**435通り**(${}_{30}C_2$)の傾きが算出され、その中央値を回帰直線の傾きとする。

外れ値の影響を受けにくい一方、データセットの真度・精度が総じて低いと、得られる中央値(すなわち求まる回帰直線の傾き)が実態に即していない可能性があることに注意が必要である。

計算が膨大なため、定められたプロセス通りに計算ができていないかの確認も必要。

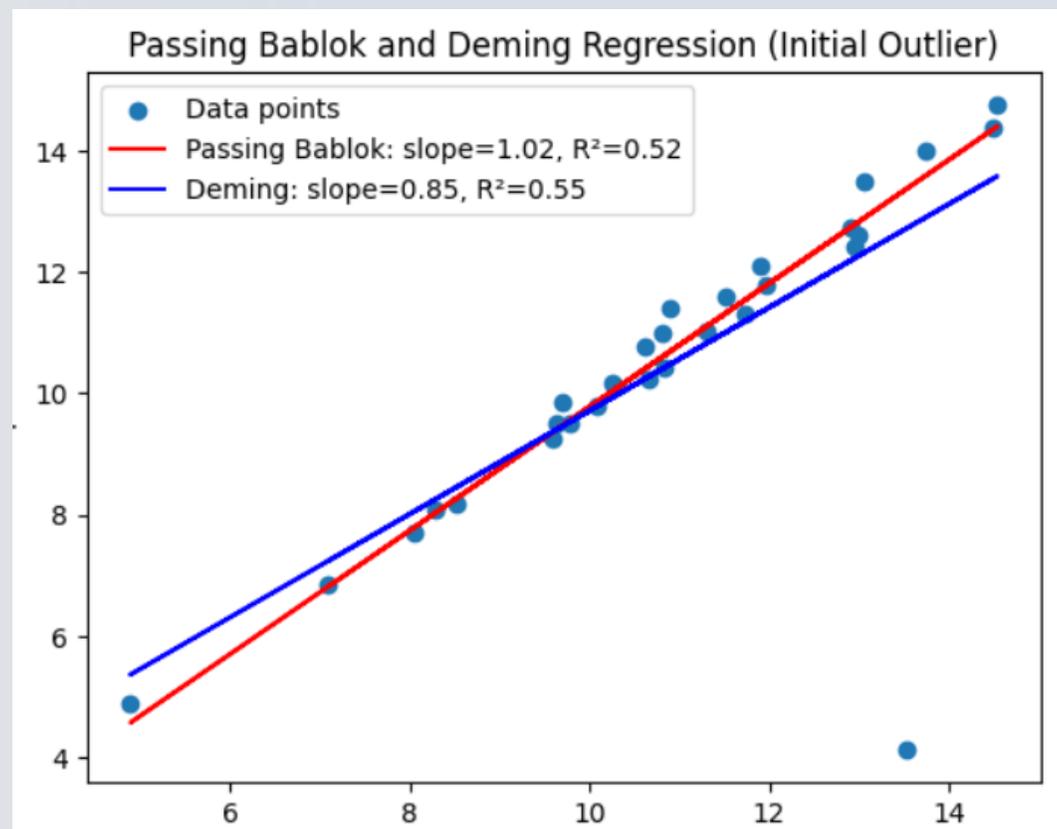
そうした措置を講じてまで本手法を使うかは関係する社内外的関係者と協議する必要がある。



Q12. Passing-Bablok回帰とDeming回帰の比較 補足

	X	Y
0	13.528105	3.851193
1	10.800314	10.824331
2	11.957476	12.555374
3	14.481786	14.048196
4	13.735116	13.702769
5	8.045444	8.276319
6	11.900177	11.933185
7	9.697286	9.669700
8	9.793562	9.803979
9	10.821197	10.808290
10	10.288087	10.320130
11	12.908547	12.935111
12	11.522075	11.498855
13	10.243350	10.264518
14	10.887726	10.872940
15	11.667349	11.665661
16	12.988158	12.983228
17	9.589683	9.568989
18	10.626135	10.624078
19	8.291809	8.276319
20	9.894020	9.893749
21	11.307237	11.308775
22	9.991095	9.990775
23	9.849008	9.849008
24	10.706573	10.706573
25	10.010500	10.010500
26	11.785870	11.785870
27	9.126912	9.126912
28	10.401989	10.401989
29	8.883151	8.883151

外れ値以外のXとYのデータの誤差は±20%程度の範囲内、
多くが±5%になるようにシミュレーション



Q13. クロスバリデーションでBland-AltmanプロットやDeming回帰を使用する場合、解釈に相違がある結果になるということはあるのか？

クロスバリデーションでBland-AltmanプロットやDeming回帰を使用する場合、一方の解析では両分析に一致性を認める結果になるが、もう一方では逆の結果になる、つまり解釈に相違がある結果になるということはあるのか？

Bland-Altmanプロットはデータのバラツキの傾向や信頼区間に入っているかどうかを確認する解析法であるのに対し、Deming回帰は両メソッドの関係性や濃度の推移におけるバイアスの変化を図示する手法である。

評価する項目と目的がそれぞれ異なるため、片方がポジティブでもう片方がネガティブと解釈される可能性もゼロではない。ただし、評価対象の両分析法はフルバリデーション済みであるため、大きく逸脱したデータが出ることは稀であると推察される。

結果に相違がある場合は、データを利用・評価する社内外の関係者とよく協議し、判断する必要がある。

Q14. クロスバリデーションではICH M10に記載の通り、サンプル数 $n \geq 30$ を採用するのが通常と考えられるが、そもそも、この値でクロスバリデーションを実施するのは妥当なのか？

例えば、ICH M10に記載のサンプル数 $n \geq 30$ という値に、どのような数値的根拠があるのかも含めて確認したい。

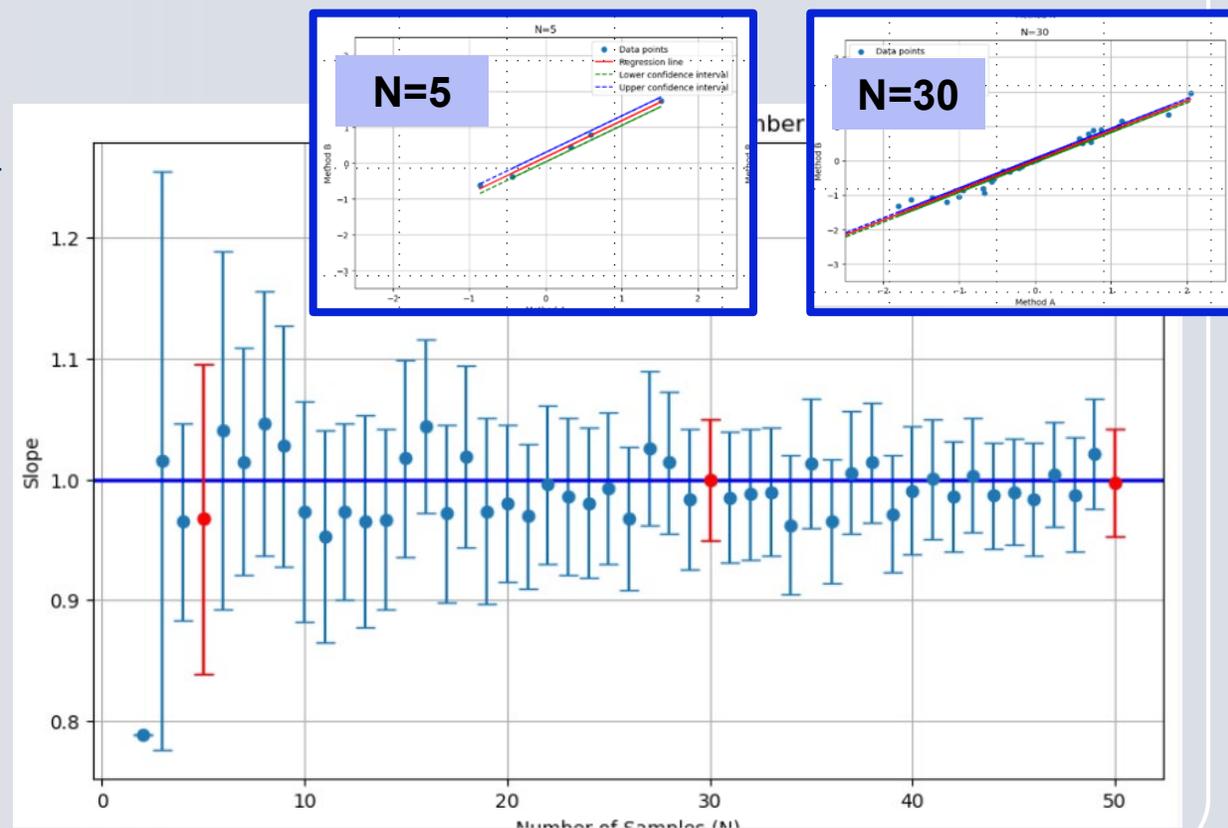
DGでは、実試料の分析数を $n \geq 30$ とする根拠となる資料を見出すことはできなかった。

しかしながら、一般的には $n=30$ の試行回数であれば、回帰モデルが十分データにフィットすると考えられるため、 $n \geq 30$ を採用したものと推察される。

右は、ある正規分布のデータセットから、 N 個のサンプルを抽出し、 N 個を分析法A(真度 $\pm 20\%$ と精度 $< 20\%$)と分析法B(真度 $\pm 20\%$ と精度 $< 20\%$)の両方で分析した線回帰時の傾きと傾きの信頼区間を示したグラフである。

※ $N=1$ から $N=50$ まで、各10回ずつ実施

$N=30$ 以上では信頼区間の幅が非常に狭いことが分かる。



The background features a light blue gradient with several 3D-rendered shapes. On the left side, there is a large, complex, organic shape in shades of blue and purple. Scattered throughout the scene are several smaller, smooth spheres of varying sizes, also in blue and purple tones, creating a sense of depth and movement.

総括と参考文献

Summary

Statistics is essential in bioanalysis to evaluate data errors and variability, leading to reliable conclusions. In bioanalysis, statistics is frequently used for assessing data reliability, creating calibration curves with weighted corrections, setting cut points for ADA analysis, and analyzing cross-validation results.

In the first half of the poster, various statistical terms related to accuracy and precision were explained step-by-step, followed by an introduction to various statistical analyses related to calibration curves and ADA analysis cut point settings. Understanding these terms and analytical methods allows for the evaluation of data reliability and the selection of appropriate statistical methods.

In the second half of the poster, the results of a survey conducted among JBF supporters were presented, addressing statistical questions and uncertainties. Additionally, some discussion points from the DG were included, explaining specific methods and approaches to deepen the understanding of statistics, along with the DG's perspectives.

参考文献①

書籍

- 丹羽 誠(2008). *これならわかる化学のための統計手法*. 化学同人.
- 芳賀 敏郎 (2011). *医薬品開発のための統計解析 第1部 基礎*. サイエンティスト社.
- 久保川 達也(2017). *現代数理統計学の基礎*. 共立出版.
- 日本統計学会(2020). *改訂版 日本統計学会公式認定 統計検定3級対応 データの分析*. 東京図書.
- 日本統計学会(2015). *改訂版 日本統計学会公式認定 統計検定2級対応 統計学基礎*. 東京図書.
- 日本統計学会(2020). *日本統計学会公式認定 統計検定準1級対応 統計学実践ワークブック*. 学術図書出版社.

ガイドライン・ガイダンス①

- International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. (2022). *ICH Harmonised Guideline: Bioanalytical Method Validation and Study Sample Analysis (M10)*.
- U.S. Food and Drug Administration. (2019). *Immunogenicity Testing of Therapeutic Protein Products — Developing and Validating Assays for Anti-Drug Antibody Detection*.
- European Medicines Agency. (2017). *Guideline on Immunogenicity Assessment of Therapeutic Proteins*.

参考文献②

ガイドライン・ガイダンス②

- MHLW(2025), ICH M10 生体試料中薬物濃度分析法バリデーション及び実試料分析 (Step 5)
- MHLW(2021), 後発医薬品の生物学的同等性試験ガイドライン(最新改訂版)
- MHLW(1996), トキシコキネティクス(毒性試験における全身的暴露の評価)に関するガイダンス(最新版)

Acceptance Criteriaに関する論文①

- Shah, V. P.; Midha, K. K.; Dighe, S.; McGilveray, I. J.; Skelly, J. P.; Yacobi, A.; Layloff, T.; Viswanathan, C. T.; Cook, C. E.; McDowall, R. D.; Pittman, K. A.; Spector, S. Analytical Methods Validation: Bioavailability, Bioequivalence and Pharmacokinetic Studies. *Pharm. Res.* 1992, 9, 588-592.
- Miller, K. J.; Bowsher, R. R.; Celniker, A.; Gibbons, J.; Gupta, S.; Lee, J. W.; Swanson, J. S. J.; Smith, W. C.; Weiner, R. S. Workshop on Bioanalytical Methods Validation for Macromolecules: Summary Report. *Pharm. Res.* 2001, 18, 1373–1383.
- Kringle, R. O. An Assessment of the 4-6-20 Rule for Acceptance of Analytical Runs in Bioavailability, Bioequivalence, and Pharmacokinetic Studies. *Pharm. Res.* 1994, 11, 556-560.
- DeSilva B , Smith W , Weiner R , et al . Recommendations for the bioanalytical method validation of ligand-binding assays to support pharmacokinetic assessments of macromolecules. *Pharm Res.*, 2003, 20, 1885-1900.

参考文献③

Acceptance Criteriaに関する論文②

- Viswanathan, C. T.; Bansal, S.; Booth, B.; DeStefano, A. J.; Rose, M. J.; Sailstad, J.; Shah, V. P.; Skelly, J. P.; Swann, P. G.; Weiner, R. Bioanalytical Method Validation—A Revisit with a Decade of Progress. *AAPS J.* 2007, 9.
- Timmerman, P.; Golob, M.; Goodman, J.; Knutsson, M.; Nelson, R.; Fjording, M. S. Toward Decision-Based Acceptance Criteria for Bioanalytical Method Validation: A Proposal for Discussion from the European Bioanalysis Forum. *Bioanalysis* 2018, 10 (16), 1255-1259.

重みづけに関する論文

- Gu, H.; Liu, G.; Wang, J.; Aubry, A.-F.; Arnold, M. E. Selecting the Correct Weighting Factors for Linear and Quadratic Calibration Curves with Least-Squares Regression Algorithm in Bioanalytical LC-MS/MS Assays and Impacts of Using Incorrect Weighting Factors on Curve Stability, Data Quality, and Assay Performance. *Anal. Chem.* 2014, 86 (18).
- Azadeh, M.; Gorovits, B.; Kamerud, J.; MacMannis, S.; Safavi, A.; Sailstad, J.; Sondag, P. Calibration Curves in Quantitative Ligand Binding Assays: Recommendations and Best Practices for Preparation, Design, and Editing of Calibration Curves. *AAPS J.* 2018, 20 (1), 22.
- Xiang, Y.; Donley, J.; Seletskaja, E.; Shingare, S.; Kamerud, J.; Gorovits, B. A Simple Approach to Determine a Curve Fitting Model with a Correct Weighting Function for Calibration Curves in Quantitative Ligand Binding Assays. *AAPS J.* 2018, 20 (1), 45.

参考文献④

ISRに関する論文

- Fast, D. M.; Kelley, M.; Viswanathan, C. T.; O'Shaughnessy, J.; King, S. P.; Chaudhary, A.; Weiner, R.; DeStefano, A. J.; Tang, D. Workshop Report and Follow-Up—AAPS Workshop on Current Topics in GLP Bioanalysis: Assay Reproducibility for Incurred Samples—Implications of Crystal City Recommendations. AAPS J. 2009, 11, 238-241.
- Hoffman, D. Statistical Considerations for Assessment of Bioanalytical Incurred Sample Reproducibility. AAPS J. 2009, 11 (3), 570.

ADA分析のカットポイント設定に関する論文

- Viswanath, G. S.; Devanarayan, V.; Amaravadi, L.; Chen, Y.; Barrett, Y.; et al. Recommendations for the Validation of Immunoassays Used for Detection of Host Antibodies against Biotechnology Products. J. Pharm. Biomed. Anal. 2008, 48 (5), 1267-1281.
- Cox, K. L.; Devanarayan, V.; Kriauciunas, A.; Manetta, J.; Montrose, C.; Sittampalam, S. Immunoassay Methods. Assay Guidance Manual 2012, 271, 143-151.

WEBページ

- 統計WEB. <https://bellcurve.jp/statistics/>
- 11th JBF Symposium DGポスター <https://bioanalysisforum.jp/common/pdf/event/dg/DG2019-43.pdf>
- 15th JBF Symposium 基礎講座 https://bioanalysisforum.jp/images/2024_15thJBFS/D1-A4-02_JBF15_Tamiki_Mori.pdf

バイオアナリシスに関する 統計学の学習法について

統計学の学習はどのようにすれば良いのか？

JBFサポーターを対象とした

当DG主催のアンケートでは、

- 統計学に関連するおすすめの書籍などを教えてほしい
- 統計学の勉強法がいまいち分からない
- 統計検定を受験する意義はあるか？

といった悩み・相談が寄せられました。



おすすめの書籍①

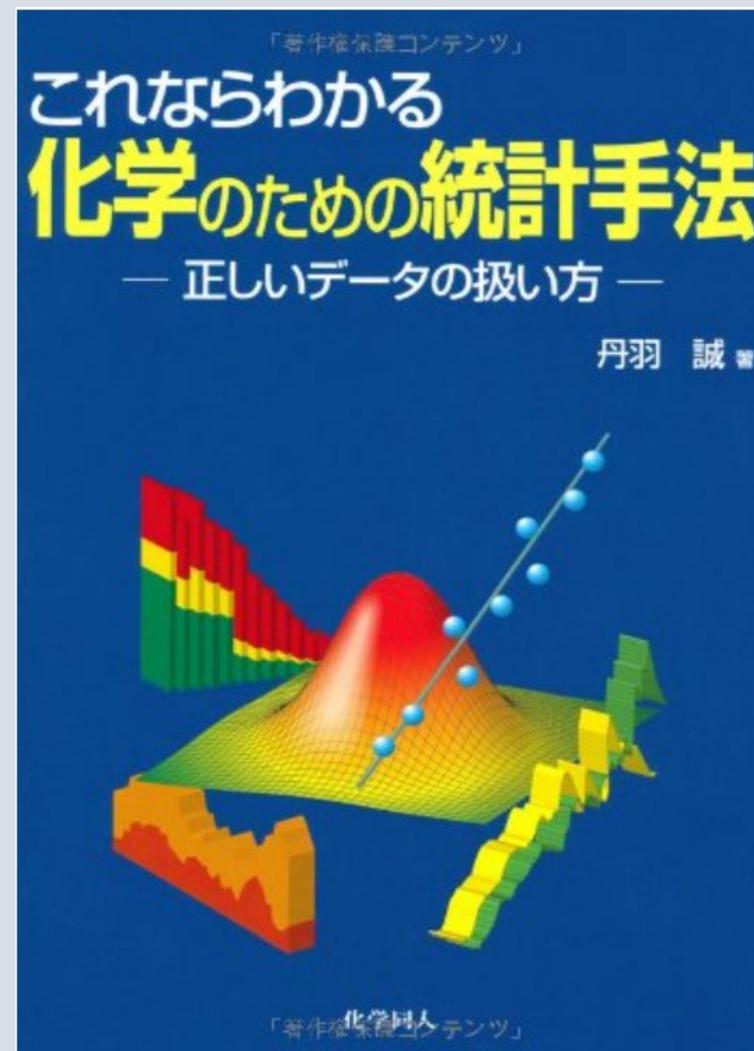
これならわかる化学のための統計手法
~正しいデータの扱い方~

丹羽誠 著 化学同人

化学実験や化学分析で使う統計処理に特化した「化学者のための」統計の書です。

図表が豊富で、バイオアナリシスで頻出するトピックが多く含まれています。

実験室で必要な統計学の基本を読み切り形式で解説しており、読みたいところだけを読むことができます。



おすすめの書籍②

医薬品開発のための統計解析

第1部 基礎 改訂版

芳賀 敏郎 著 サイエントリスト社

統計ソフト JMP と Excel が搭載されている PCを使って、医薬品開発に関する統計解析を基礎から学習できるようになっています。

ただ読むだけでなく、ダミーデータを使用して解析をする練習するため、より実践的に統計手法を理解できます。

式の導出やグラフなども非常にわかりやすく、非常に丁寧な解説です。



模擬データは以下よりダウンロード可能
<https://scientist-press.com/download/>

医薬品開発のための統計解析（通称グリーン本） を用いた勉強会について

【日程】 2025年5月スタート 毎月2回（原則第1, 3金曜日）の頻度で開催
※詳細な日程は Excel で適宜案内予定

【時間帯】 16:00-18:00

【開催場所】 オンライン開催（Zoom）

【準備するもの（勉強会そのものは費用無料）】

①Excel がインストールされた PC

※JMP（Ver.9 以降）のインストールを推奨。

大学等の所属の場合、JMPはStudent Editionが無料入手可能です。

②芳賀敏郎著「医薬品開発のための統計解析第1部基礎 改訂版」

※あらかじめ購入していただく

③オンライン講義が受講可能な PC 環境

【連絡先】 福島 慎二さん s.fukushima.stat@gmail.com

参加者の氏名およびメールアドレスはメンバー全員に開示していただきます

おすすめのサイト

統計WEB ~統計学の時間~

統計学の基礎から応用までを丁寧に解説しています。

具体例や練習問題を多数掲載し、初心者でも理解しやすい内容です。

統計検定®2級の範囲をカバーし、学習ページ、練習問題、解説を通じて、統計学の知識を深めることができます。



統計学の基礎から応用までを丁寧に解説しています。「Step1. 基礎編」は、大学で学ぶ統計学の基礎レベルである統計検定®2級の範囲をほぼ全てカバーする内容となっています。最後まで読み進めることで、統計検定®2級に合格できる力がつくことを目標としています。

<https://bellcurve.jp/statistics/course/?srsId=AfmBOoqSeywG6GTg8yqa1xGxXCa6TNoFBmsmlvx70tIkuVGulaSmqvQJ>

統計学の勉強方法や

統計検定の受験の意義は？

統計学を体系的に勉強したい場合は、前述の書籍や論文を読むことの他に統計検定がおすすめです。ただし、**バイオアナリシスに関連のない単元も多く含むことにご留意ください。**

統計学の知識	統計検定の級数	バイオアナリシスとの関連
真度・精度関連の用語 (バイアスや標準偏差、変動係数等)	3級	Acceptance Criteria
各種グラフの理解 (ヒストグラム・箱ひげ図・散布図等)	3級	分析結果のグラフ化 外れ値解析
回帰分析全般 (R^2 値、相関係数など)	3～2級	検量線 クロスバリデーション
t分布 統計的仮説検定 ノンパラメトリック解析 分散分析(ANOVA)	→3～2級 →3～2級 →単元になし →2～準1級	ADA分析のカットポイント設定