

ISO/TC69（統計的手法の適用）で 開発された不確かさに関わる規格

（独）産業技術総合研究所
計測標準総合センター
小池昌義

目次

- ISO/TC69 統計的方法の適用
- TC69と不確かさ関連JIS規格の制定
- JIS Z 8404-1 (ISO/TS 21748:2004)
- JIS Z 8404-2 (ISO/TS 21749:2004)
- まとめ

ISO/TC69 統計的方法の適用

WG 3: Statistical interpretation of data

WG 9: Random sampling generation

Ad Hoc's Evaluation of conformity/ Use of software
/Application of ISO/IEC 17025

SC 1: Terminology and symbols(3534)

SC 4: Process management

SC 5: Acceptance sampling

SC 6: Measurement methods and results

SC 7: Implementation of six sigma

SC 8: New technology and product development

TC69/SC6

ISO/TC69/SC6 “Measurement methods and results”

WG1: Accuracy of measurement methods and results
(5725)

**WG4: Statistical aspects of
the preparation and use of reference materials**

WG5: Capability of detection

**WG7: Statistical methods
to support measurement uncertainty evaluation**

ISO/TC69と不確かさ関連規格

ISO/TC69 統計的方法の適用

SC6 測定方法とその結果

WG7 不確かさのための統計的方法

SC1 & SC6

ISO 3534 シリーズ
ISO 5725 シリーズ

JIS Z 8101 シリーズ
JIS Z 8402 シリーズ

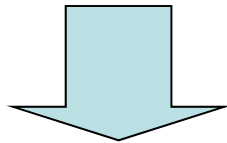
ISO/TS 21748:2004
ISO/TS 21749:2004

JIS Z 8404 測定の不確かさ

JIS Z 8404-1
JIS Z 8404-2

ISO/TC69の開発した規格(例)

ISO 3534シリーズ 用語と記号



JIS Z 8101 統計—用語と記号—

- 1, —第1部: 確率及び一般統計用語
- 2, —第2部: 統計的品質管理用語
- 3, —第3部: 実験計画法

ISO/TC69の開発した規格(例)

ISO 5725シリーズ

測定方法及び測定結果の精確さ(真度及び精度)

JIS Z 8402-1 — 第1部: 一般的な原理及び定義

JIS Z 8402-2 — 第2部: 標準測定方法の**併行精度**及び
再現精度を求めるための基本的な方法

JIS Z 8402-3 — 第3部: 標準測定方法の**中間精度**

JIS Z 8402-4 — 第4部: 標準測定方法の**真度**を
求めるための基本的な方法

JIS Z 8402-5 — 第5部: 標準測定法の**精度**を
求めるための**代替法**

JIS Z 8402-6 — 第6部: **精確さ**に関する値の実用的な使い方

JIS 8101/JIS 8103

統計用語と計測用語の対比

概念

ばらつきの小さい程度

かたよりの小さい程度

総合的な良さ

短時間の間の

複数の測定 (repetition)

条件を変えた

複数の測定 (replication)

統計用語／計測用語

精度／精密度 (精密さ)

真度／正確度 (正確さ)

精確度 (精確さ)／精度

併行測定／繰返し測定

繰返しまたは反復測定／
反復測定

JIS 8101/JIS 8103

統計用語と計測用語の対比

英語：統計用語／計測用語

precision : 精度／精密度

trueness : 真度／正確度

accuracy : 精確さ／精度

repeatability

併行精度／繰返し性

repeatability condition

併行条件／繰返し条件

intermediate precision

中間精度／（再現性）

reproducibility

再現精度／再現性

reproducibility condition

再現条件／（再現条件）

「中間精度」とは、

4つの条件（時間、校正、オペレータ、使用装置）のうちいくつか
が異なる条件の下での「精度」である。

JIS Z 8404-2(ISO/TS 21749:2004)

ISO/TS 21749

**(Measurement Uncertainty for metrological application
– Repeated measurements and nested experiments)**

**測定の不確かさ—第2部：測定の不確かさの評価における
繰返し測定及び枝分かれ実験の利用の指針**

JIS Z 8404-2 目次

序文／1 適用範囲 ／2 引用規格 ／3 用語及び定義3

4 不確かさの評価の統計的方法

4.1 計測における不確かさの表現のガイドのアプローチ

4.2 チェック用標準

4.3 不確かさの評価のステップ

4.4 この規格における例

5 不確かさのタイプA評価

5.1 一般

5.2 不確かさのタイプA評価における時間の役割

5.3 測定方法

5.4 材料の不均質性

5.5 測定の方法によるかたより

6 タイプBの不確かさの評価方法

7 不確かさの伝ば(播)

7.1 一般 ／7.2 一変数関数に関する公式 ／7.3 二変数関数に関する公式

8 事例－ゲージ調査からのタイプAの不確かさの評価方法

8.1 目的及び背景

8.2 データの収集及びチェック用標準

8.3 併行精度, 日間及び長時間効果の解析

8.4 プローブのかたより ／8.5 配線のかたより

8.6 不確かさの計算

附属書A(規定)記号 ／附属書JA(参考)参考文献

JIS Z 8404-2(ISO/TS 21749:2004)

測定の不確かさ—第2部：測定の不確かさの評価における 繰返し測定及び枝分かれ実験の利用の指針

1 適用範囲

この規格は、GUMに採用されているアプローチに従って、不確かさの成分を求め、合成する基本的な手順を確立する。この基本手順には、個々の成分、特にタイプA評価の不確かさとして分類される、統計的方法に基づく不確かさ成分を評価するために分散分析(ANOVA)を使用する統計的枠組みが追加されている。

この規格は、不確かさの成分が、繰返し測定、測定機器、試験項目又は検査用標準の統計解析から見積もることができるような実験的状况を規定する。

この規格は、1段、2段又は3段の枝分かれ実験から不確かさを見積もる場合だけを規定する。より複雑な実験的状况は取り扱わない。

不確かさの評価で従うべきステップの概要

a) タイプAの評価方法:

1) 出力量を Y で表し、 Y を繰返し測定できる場合は、次の偶然効果の分散成分の推定値をもとめるためにANOVAモデルを使用する。

- 試験品目に対する繰返し測定の結果のばらつき
- チェック用標準の測定のばらつき
- 実験計画に従って行った測定のばらつき

2) Y を直接繰返し測定できず、また、次のモデル

$$Y = f(X_1, X_2, \dots, X_n)$$

が知られていて、

更に、入力量 X_i を繰返し測定できる場合は、

X_i の最適推定値 x_i の不確かさを評価し、不確かさの伝ば(播)則を使用できる。

3) Y 又は X_i の測定を繰り返すできない場合は、タイプBの評価方法を参照する。

b) タイプBの評価方法: 各入力量の最適推定値の標準不確かさを評価する。

c) 測定結果の標準不確かさを得るために、

タイプA及びタイプBの評価方法からの標準不確かさを合成する。

例：製造製品検査の測定データ

測定プロセスの効果
(測定の変動)

+

製造プロセスの効果
(製造の変動)

チェック用標準の測定データの利用により分離する

チェック用標準とは、次の特性をもつことが必要とされる標準である。

- a) 定期的に測定できなければならない。
- b) 生産品目に対して、組成及び形状が近似していなければならない。
- c) 安定した人工物でなければならない。
- d) 常に、測定プロセスで利用可能でなければならない。

時間に依存した効果の変動の推定が可能

- a) 短期的変動(併行精度又は測定機器精度)
- b) 中間的変動(中間精度として知られる。)
- c) 長期的変動[ラン間の変動, 又は安定性]

分散分析の活用

枝分かれ実験計画(nested design):

ある因子のすべての水準が、他のすべての因子の一つの水準だけに現れる実験の計画。

段階1: 測定の併行精度を得るために、
短時間に行われた測定

段階2: 数日間を経た後、日を変えて
(又は適切な時間増分後)の測定

段階3: 月ごとに分けて行われた一連の測定

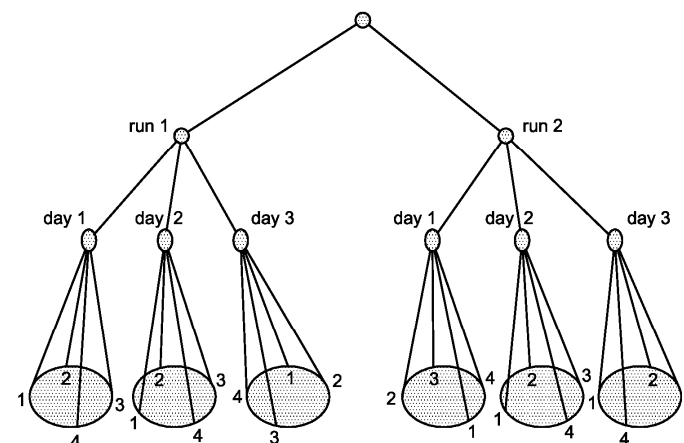


表 1-3 段枝分かれ計画の分散分析表

要因	自由度 v	平方和 SS	平均平方 MS	平均平方の期待値
ラン	$L - 1$	SS_R	MS_R	$\sigma^2 + J\sigma_D^2 + JK\sigma_R^2$
日 (ラン内)	$L(K - 1)$	$SS_{D(R)}$	$MS_{D(R)}$	$\sigma^2 + J\sigma_D^2$
誤差	$LK(J - 1)$	SS_E	MS_E	σ^2

変動の要因、平方和 (SS) 及び対応する自由度 (v) を、それぞれ、第 1 列、第 3 列及び第 2 列に示す。平方和を対応する自由度で除した平均平方 (MS) は、第 4 列に示してある。最終列は、平均平方の期待値である。

Gauge study(1)

シリコンウェハの抵抗率の測定:

不確かさの評価では、次の時間に依存したばらつきの原因を考慮した。

- a) ウェハ中心における測定の繰返し
- b) 日間効果
- c) ラン間の効果

さらに、次のかたよりの原因を考慮した。

- プローブ No.2362 によるかたより
- 配線の方法 A によるかたより

3段枝分かれ計画から推定した。

- $J = 6$ 個のウェハ中心における測定値
- $K = 6$ 日間
- $L = 2$ のラン

3段枝分かれ計画に対するモデルは、次のとおりである。

$$Y_{lkj} = \mu + \gamma_l + \delta_{lk} + \varepsilon_{lkj}$$

Gauge study (2)

表 9—ゲージ調査データの分散分析表

要因	自由度	平均平方	平均平方の期待値
ラン	1	$MS_R = 0.009\ 198$	$\sigma^2 + 5\sigma_D^2 + 30\sigma_R^2$
日 (ラン内)	10	$MS_{D(R)} = 0.003\ 238$	$\sigma^2 + 5\sigma_D^2$
誤差	44	$MS_E = 0.000\ 804\ 6$	σ^2
注記 この例において、クロスオーバ効果（ウエハと実験との交互作用効果）は、考慮していない。こうした効果は、一つの要因が他の原因に影響を及ぼす場合に考慮する。例えば、一方の薬剤の投与が他方の効果を抑制するような状況で、2種類の薬剤の効果を調査する場合が該当する。			

併行精度、日間及び繰返し効果による分散成分も得られる。

$$\hat{\sigma}^2 = S^2 = MS_E = 0.000\ 804\ 6$$

$$\hat{\sigma}_D^2 = S_D^2 = \frac{MS_{D(R)} - MS_E}{5} = 0.000\ 486\ 6$$

$$\hat{\sigma}_R^2 = S_R^2 = \frac{MS_R - MS_{D(R)}}{30} = 0.000\ 198\ 7$$

JIS Z 8404-1(ISO/TS 21748:2004)

ISO/TS 21748

(Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation)

測定の不確かさー

第1部：測定の不確かさの推定における併行精度，
再現精度及び真度の推定値の利用の指針

計測用語で表現すれば、

「測定の不確かさの評価における

繰返し性、再現性及び正確さの推定値の利用の指針」

JIS Z 8404-1(ISO/TS 21748:2004)

測定の不確かさ－第1部：測定の不確かさの推定における 併行精度，再現精度及び真度の推定値の利用の指針

1)適用範囲 この規格は，次のための指針である

－JIS Z 8402-2に従って実施した(共同)実験から得られた値を利用した
測定の不確かさの評価。

－共同実験による結果と不確かさの伝播則を用いて得られた測定の不確かさ(MU)との比較(14章参照)。

引用規格

JIS Q 0033 認証標準物質の使用_(ISO Guide 33)

JIS Z 8402-2, 測定方法及び測定結果の精確さ(真度及び精度)－第2部：
標準測定方法の併行精度及び再現精度を求めるための基本的方法 (**ISO 5725-2**)

JIS Z 8402-3 測定の方法及び測定結果の精確さ(真度及び精度)－第3部：
標準測定方法の中間精度 (**ISO 5725-3**)

JIS Z 8404-1 目次

序文／1. 適用範囲／2. 引用規格／3. 定義／4. 記号

5. 原理

5.1 個々の測定結果と測定プロセスのパフォーマンス

5.2 再現精度の値の適用性

5.3 統計的モデルの基本式

5.4 併行精度の値

6. 併行精度, 再現精度及び真度の推定値を用いた不確かさの評価

6.1 測定の不確かさの評価の手順

6.2 予想される精度と現実の精度の差

7. 測定方法のパフォーマンス値と,

特定の測定プロセスから得られる測定結果との関連性の確認

7.1 一般

7.2 かたよりの試験所成分の管理状態の実証

7.3 併行精度の確認

7.4 パフォーマンスの継続的検証

8. 試験品との関連性の確認

8.1 一般

8.2 サンプルング

8.3 試料の調製及び前処理

8.4 試験品の種類の変更

8.5 応答水準による不確かさの変化

9. 付加的な要因

10. 合成標準不確かさの一般的表現

11. 共同実験結果に基づく不確かさのバジェット表

12. 合成された結果に対する不確かさの評価

13. 不確かさの情報の表現

13.1 一般的表現

13.2 包含係数の選択

14. 方法のパフォーマンス値と不確かさの値の比較

14.1 比較のための基本となる仮定

14.2 比較手順

14.3 差異の原因

附属書A(参考)不確かさの二つの評価方法

A.1 GUMの方法

A.2 共同実験方法

A.3 方法間の関係

附属書B(参考)実験による不確かさの評価

B.1 感度係数推定のための実践的手順

B.2 ランダムな効果による不確かさを評価するための簡単な手順

附属書C(参考)不確かさの計算例

C.1 自動車排気ガス中の一酸化炭素(CO)の測定

C.2 食肉成分の定量

参考文献

共同実験の結果の利用

JIS Z 8401-2に基づいた共同実験:

試験方法の

真度に関する推定値 s_L

併行精度に関する推定値 s_r

再現精度に関する推定値 s_R

試験所: 測定プロセスの管理と保証の裏付け

- ・かたよりの試験所成分が管理状態にある
- ・併行精度(ばらつき)が管理状態にある
- ・継続的なパフォーマンスチェック

適用可能

共同実験で取り扱われなかった成分の確認

方法の真度に関する不確かさの推定値
再現精度の推定値
付加的な成分の影響

合成標準不確かさの評価

6.1 併行精度，再現精度及び真度の推定値を用いた 不確かさの評価（手順）

- a) 公開されている情報から，併行精度，再現精度，及び真度の推定値を入手する。
- b) 測定値に対する試験所のかたよりが、推定値から予想される範囲内に
あるか否かを確認する。（管理状態の確認）
- c) 現在の測定で実現されている併行精度及び再現精度が、推定値から予想される
範囲内にあるか否かを確認する。（管理状態の確認）
- d) 6.1 a)における検討では適切に取り入れられていないあらゆる効果に
よって生じ得るばらつきを定量化する。（その他の効果の見積もり）
- e) かたより及び精度が管理状態にある場合は，
再現精度の算定値[6.1 a)]と，真度に付随する不確かさ[6.1 a)及びb)]，
及び付加的な成分の影響 [6.1 d)]を合成して，
合成不確かさを評価する。

7. 測定方法のパフォーマンス値と、特定の測定プロセスから得られる測定結果との関連性の確認

7.1 一般 共同実験の結果から、パフォーマンスの指標(s_R , s_r) が得られ、ある場合には、測定方法のかたよりの推定値が得られる。これらは測定方法のパフォーマンスに関する“仕様”を構成する。特定の目的のための測定方法を採用するに当たり、試験所は通常、この“仕様”が満たされていることを確認することが期待される。

7.2 かたよりの試験所成分の管理状態の実証

7.3 併行精度の確認

- 共同実験で得られた併行標準偏差と整合していること

7.4 パフォーマンスの継続的検証

- かたより及び精度の定期的チェックを含む適切な測定の質の管理。
- 測定の質の保証手段。

7.2 かたよりの試験所成分の管理状態の実証

7.2.2 かたよりの試験所成分が管理状態にあることを実証する方法

7.2.2.2 認証標準物質又は参照標準の利用

7.2.2.3 不確かさが既知の基準試験方法との比較

7.2.2.4 同じ方法を使用する他の試験所との比較

**7.2.3 有意なかたよりの試験所成分が検出された場合
適用範囲外である。
かたよりの原因を調査し、取り除くこと。**

8. 試験品との関連性の確認

8.1 一般

8.2 サンプルング

8.2.1 サンプルングプロセス

8.2.2 不均質性

8.3 試料の調製及び前処理

8.4 試験品の種類の変更

8.5 応答水準による不確かさの変化

8.5.1 s_R の調整 (JIS Z 8402-2 参照)

8.5.2 不確かさに対する他の寄与の変化

10. 統計的モデルの基本式

9. 共同実験で考慮されなかった成分

観測結果の
統計的モデル

$$y = \mu + \delta + B + \sum c_i x'_i + e \quad \leftarrow e : \text{誤差項}$$

B : かたよりの試験所成分

δ : 測定方法に固有のかたより

不確かさのモデル

$$u^2(y) = u^2(\hat{\delta}) + s_L^2 + \sum c_i^2 u^2(x_i) + s_r^2$$

$$u^2(y) = u^2(\hat{\delta}) + s_R^2 + \sum c_i^2 u^2(x_i) \quad (12)$$

δ の不確かさ

$$u(\hat{\delta}) = s_{\hat{\delta}} = \sqrt{\frac{s_R^2 - (1 - 1/n)s_r^2}{p}}$$

s_R : 再現標準偏差 $s_R^2 = s_L^2 + s_r^2$

11. 共同実験に基づく不確かさのバジェット表

効果	y に関連する標準不確かさ ^(a)	コメント
δ	$u(\hat{\delta})$	かたよりに対する補正が共同実験に組み込まれ, かつ不確かさが無視できない場合にだけ含まれる。
B	s_L	表 2 参照。
e_r	s_r	方法 ^(b) の n_r 回の反復の平均値が試験品に対して実際に使われた場合, e_r に付随する不確かさは $s_r/\sqrt{n_r}$ となる
x_i	$c_i u(x_i)$	8 章及び 附属書 B を参照。
注^(a) これらの標準不確かさの単位は y と同じである。 (b) 方法それ自体が, 反復を指示する場合がある。 n_r は, こうした反復を含めた方法全体の繰返しに関係がある。		

13. 不確かさの情報の表現

13.1 一般的表現 不確かさは、合成標準不確かさ又は拡張不確かさ

13.2 包含係数の選択

13.2.1 一般 包含係数を選択するにあたって次の事項を考慮することが重要である。

13.2.2 望ましい信頼の水準 約95%の信頼の水準に対応。目的による。

13.2.3 推定値の自由度

13.2.3.1 約95 %の信頼の水準で、支配的な寄与の自由度が大きい(>10)場合、ほとんどの現実的な目的では、 $k=2$ を選択してもよい。

13.2.3.2 一つの項が支配的である場合[表示レベルが $u_i(y) \geq 0.7u(y)$ の場合]、通常は、の有効自由度はとすれば十分である。

13.2.3.3 複数の重要な項がほぼ同じ大きさを、限られた自由度である場合は、ウェルチ=サタスウェイトの式を適用する。

備考 測定及び試験の多くの分野では、統計的外れ値の頻度が、正規分布からの期待値と比べてかなり高いために、当該の分布についての十分な知識をもたずに高い信頼の水準(>95 %)を推定するにあたっては、きわめて慎重な配慮が必要である。

14.方法のパフォーマンス値と不確かさの値の比較

14.1 比較のための基本となる仮定（通常の検定の利用）

第一に、**標準不確かさが表す分布は、正規分布である**という仮定。

しかし、様々な分布からの様々な大きさの分散をもつ項に付随する不確かさがあるから、検定は示唆的なものとして取り扱い、信頼の水準は十分慎重に見ることが望ましい。

第二に、**比較される二つの不確かさの評価値は、完全に独立している**という仮定である。

しかし、実際には、いくつかの要因は両方の評価値に共通している。

二つの不確かさの評価値に、重要な要因が共通する場合、

二つの評価値は偶然によって類似するよりもはるかに高い頻度で類似するであろう。

このような場合、次に示す検定で有意な差異が発見できないからといって、

その結果は測定モデルの信頼性を示す強力な証拠と見なすべきではない。

14.2 比較手順 F検定により、分散の差異を比較する。

14.3 差異の原因 差異の理由には、次のことが含まれる。

- －試験所間に存在する能力の真の違い
- －測定に及ぼす重要な効果のうち、モデルに含まれていないものがある
- －不確かさに対する重要な寄与の過大評価又は過小評価

附属書A(参考)不確かさの二つの評価方法

- GUMの方法による評価：
個々の入力量を重視→ボトムアップの方法
- 5725-2共同実験の結果に基づく評価：
方法全体のパフォーマンスを重視→トップダウンの方法
- その他の影響－合成モデル：
ハイブリッドアプローチ

2つの異なる推定値の比較は、
測定モデルの完全性の評価として有効である。

私見：不確かさの評価方法：A and B

タイプAの評価方法

「一連の観測値の統計的解析による不確かさの評価の方法」、

タイプBの評価方法

「一連の観測値の統計的解析以外の手段による不確かさの評価の方法」

上の定義での、「一連の観測値」の源 (A or B) と

「解析手段」の分類 (A' or B') と

を区別して分類をすれば、以下のように4通りになる。

AA' : その測定データの、統計的解析による評価

BA' : 過去のデータの、統計的解析による評価

AB' ; その測定データの、統計的解析以外の手段による評価

BB' ; 過去のデータの、統計的解析以外の手段による評価

このような分類を考えれば、この規格の方法はBA'に対応する。

私見：不確かさの評価方法：A and B

その場での複数の観測データ

同じ条件での過去の観測データ

同じ方法の共同実験データ

その他の実験データ

繰返し数が少ないと、信頼性がない。
繰返し数が大きいと、コストが増大。

**データ数を増し、
信頼性を上げる。**

過去のデータが利用できる条件を
検証することが必要である。
そのためには、
統計的方法の正しい適用が必要。

実際のデータに基づいた、信頼性の大きい標準不確かさを見積もる。

私見：不確かさの評価方法：検証例

その場での複数の観測データ

← DUTによる成分を反復・繰返しの標準偏差 σ 又は区間Rを用いて検証する。

同じ条件での過去の観測データ

← DUT以外の不確かさ成分は同じ手順、同じ条件であり、利用可能。

同じ方法の共同実験データ

← JIS Z 8404-1の方法を利用。

その他の実験データ

← 実験計画に基づいたデータであれば、データの素性が明確であり、利用可能。

**統計的方法の正しい適用が
信頼性の高い標準不確かさ評価には必要。**

不確かさ・統計的方法に関わる他の規格

- ・ JIS Z 8405 試験所間比較による技能試験のための統計的方法
- ・ JIS Z 8403 製品の品質特性－規格値の決め方通則
- ・ JIS Z 9090 測定－校正方式通則(附属書)
- ・ ISO/IEC G98－3 (GUM): 近日発刊、JIS化
- ・ その他、統計的方法の規格は多数。
- ・ 規格に頼らず、統計の教科書は多数。

まとめ

- ISO/TC69は「統計的方法の適用」の立場から多くの規格を発行してきた。
- GUM発行以前の規格も含めて、測定の不確かさの評価において利用可能である。
- データの統計的な評価を行った結果を不確かさ評価に利用する場合、どのような状況で評価された結果であるかをみて、利用できる範囲と条件を見極めることが必要である。
- データに基づいた、信頼性の高い標準不確かさの評価には、統計的方法の活用がキーである。