不確かさガイド(GUM)の改訂動向とベイズ統計の利用

産業技術総合研究所　計量標準総合センター
榎原研正
ehara.kensei@aist.go.jp

Outline

1. GUM改訂の動き
2. ベイズ統計とは
3. 改訂案の概要
4. 現状整理

*) GUM: Guide to the Expression of Uncertainty in Measurement（測定における不確かさの表現のガイド), ISO (1995) ≅ JCGM 100 (2008)

GUMの発展

- 1980 Recommendation INC-1
- 1993 GUM 第1版
- 1995 〃 (微修正)
- 2008 JCGM 100 (追加微修正, WEB上でオープン化: www.bipm.net/en/publications/guides/gum.html)

- ✓ GUM第１版から内容は実質上変わっていない
- ✓ 基礎科学、計量標準の国際比較、試験・校正機関の認定、トレービリティの保証など、様々な局面で利用されるようになっている

GUMのメンテナンス機関

JCGM (Joint Committee for Guides in Metrology)

= 国際度量衡局(BIPM)に事務局をおき、次の8機関をメンバー機関とする国際委員会

- BIPM: Bureau International des Poids et Mesures
- IEC: International Electrotechnical Commission
- IFCC: International Federation of Clinical Chemistry and Laboratory Medicine
- ILAC: International Laboratory Accreditation Cooperation
- ISO: International Organization for Standardization
- IUPAC: International Union for Pure and Applied Chemistry
- IUPAP: International Union for Pure and Applied Physics
- OIML: International Organization of Legal Metrology

❑ JCGM-WG1: GUM担当

❑ JCGM-WG2: VIM担当

日本からは今井秀孝氏(NMIJ客員研究員, NITE客員調査員, 元産総研理事)が参加(ILAC代表)

GUMの周辺文書

- 国際計量計測用語
  - JCGM 200 [= VIM第3版] (2007)

ベイズ統計に配慮

- GUM補完文書
  - JCGM 101 [= GUM Supplement 1] (2008)
    "モンテカルロ法を用いた分布の伝播"
  - JCGM 102 [= GUM Supplement 2] (2011)
    "出力量が複数ある場合への拡張"
  - (審議中) JCGM 103 (= GUM Supplement 3) "モデリング"

ベイズ統計に沿う

ベイズ統計に沿う

- その他のJCGM文書
  - JCGM 104 (2009) "GUMとその関連文書の紹介"
  - JCGM 106 (2012) "適合性評価における測定不確かさの役割"
  - (審議中) JCGM 105 "測定不確かさの評価のための概念,原理,方法"
  - (審議中) JCGM 107 "最小二乗法の応用"

ベイズ統計に沿う

GUM改訂に向けた動き (1)

周辺文書へのBayes統計の導入

- 2007: VIM(国際計量基本用語集)第3版
- 2008: GUM Supplement 1 (モンテカルロ法を用いた分布の伝播)
- 2011: GUM Supplement 2 (出力量が複数ある場合への拡張)
- 2012: JCGM 106(適合性評価における測定不確かさの役割)

解説論文等による事前説明

- 2008: "How to revise the GUM?" (W. Bich)
- 2012: "Revision of the Guide to the Expression of Uncertainty in Measurement" (W. Bich他)

GUM改訂に向けた動き (2)

事前アンケート調査

- 2012実施，改訂の必要性, Bayes統計利用の賛否など
- 回答の集計・解析がBIPMホームページで公開されている

改訂案執筆

- 200? - 2014/12月，JCGM-WG1

改訂案コメント収集

- 2014/12月- 2015/4月
- JCGMメンバー機関(MOs) + 国家計量標準機関(NMIs)

ワークショップ

- 2015/6月： BIPM Workshop on Measurement Uncertainty

(www.bipm.org/en/conference-centre/bipm-workshops/measurement-uncertainty/)

コメント集計（C. Michotte, BIPM）

		Member Organizations: 6 MOs	National Metrology Institutes: 25 NMIs	合計
本体	JCGM 100 (GUM rev.)	530	375	905
事例集	JCGM 110 (exples)	83	85	168
				1073

・・・ 改訂に否定的な意見が優勢であった

ISO/IEC ballots: « Do you agree to the circulation of the drafts as a DIS? »

	Yes	Yes+comments	No	Abstain	Result
ISO	9	1	2	3	✓
IEC / JCGM 100	6	4	3	1	✓
IEC / JCGM 110	6	4	2	2	✓

GUM改訂案

2分冊とした（本体＋事例集）

- 本体：JCGM 100　201X CD
  Evaluation of measurement data — Guide to Uncertainty in Measurement （測定データの評価 — 測定における不確かさのガイド)

- 事例集：JCGM 110　201X CD
  Evaluation of measurement data — Examples of uncertainty evaluation (測定データの評価 — 不確かさ評価の事例)

✓　公開文書ではない (CD = Committee Draft)

GUM改訂案本体目次

1. Introduction （序）
2. Scope （適用範囲）
3. Summary of procedure for evaluating and reporting uncertainty （不確かさの評価と報告の手順のまとめ）
4. Normative references （引用文献）
5. Terms and definitions （用語と定義）
6. Conventions and notation （慣用と記号）
7. Basic concepts （基本概念）

現行GUMとは構成がやや異なる

8. Modelling the measurement （測定のモデル化）
9. Estimating the input quantities and evaluating standard uncertainties and covariances （入力量の推定と標準不確かさ及び共分散の評価）
10. Estimating the measurand and evaluating the standard uncertainty （測定対象量の推定と標準不確かさの評価）
11. Determining a coverage interval for the measurand （測定対象量の包含区間の決定）
12. Reporting and recording measurement results （測定結果の報告と記録）

GUM改訂案事例集目次

1. Gauge block calibration （ブロックゲージの校正）
2. Simultaneous measurement of resistance and reactance （電気抵抗及びリアクタンスの同時測定）
3. Measurement of Celsius temperature using a resistance thermometer（抵抗温度計を用いた温度測定）
4. Molecular weights （分子量）
5. Gravimetric mixture preparation and the calculation of molar composition （質量比混合法とモル組成の計算）

- ✓ 現行GUMから「放射能濃度測定」,「分散分析」,「硬さ測定」が消え，4.と5.の事例が追加された
- ✓ 事例を別冊にすることで随時他の事例も追加，との趣旨

改訂案ラフスケッチ(1) 評価手続き上の変更点

- タイプA評価 (繰り返し数 n)

現行 $\frac{s}{\sqrt{n}}$ ⇨ 改訂案 $\sqrt{\frac{n-1}{n-3}}\frac{s}{\sqrt{n}}$ （大きくなる）

- タイプB評価 ・・・ （多くの場合）現行と同じ
- 不確かさ伝播則 ・・・ 現行と同じ
- 拡張不確かさ (包含区間)

現行

t 分布表に基づく包含係数 k
(多くの場合 $k \fallingdotseq 2$ [$\nu_{\mathrm{eff}} = \infty$])

⇨

改訂案 以下から選択:

(1) 分布形非依存の k [=2.98(対称分布), =4.47(非対称分布) for 95%包含確率]

(2) 確率分布から k を決める
  (2-a) 正規分布を想定できる場合、$k = 1.96$
  (2-b) それができない場合、確率分布を数値的に求める(Monte Carlo法)

(3) 解析的に確率分布を求める(特殊ケース)

改訂案ラフスケッチ(2) 概念上の変更点

- ベイズ統計の利用
- (有効)自由度の概念は消滅
- 用語の定義から、タイプA評価、タイプB評価、拡張不確かさなどが消滅 (±拡張不確かさ → 包含区間)
  [これらの用語をなくすわけではない。重要度が下がった]

GUM改訂の背景 (論文2等に基づく)

改訂の動機

現行GUMには

- 内部的不整合がある(Type AとType Bで確率の意味が異なる)
- 外部的不整合がある(GUMと、ベイズ統計を利用するSupplement 1 & 2、及びVIM3が不整合)

基本方針

- ベイズ統計の採用により、これらの不整合を解消する
- 内容の難易度を維持する

ベイズ統計とは

1. ベイズの定理を(必ず)用いる
2. 確率の概念を広く捉える

(例) ・基礎物理定数を確率変数とみなす

・過去のある時点・ある場所の天気を確率変数とみなす

→ 主観確率 (“state of knowledge” を表す)

<u>頻度主義的統計学(伝統的統計学)では・・・</u>

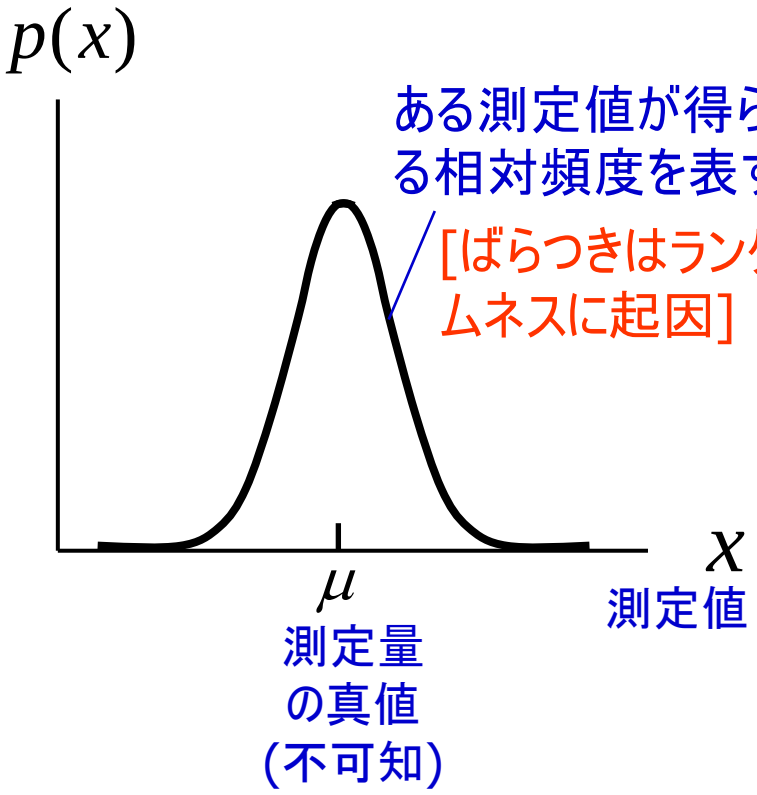
- ある試行を n 回繰り返したとき，ある事象が r 回生じた

  ⇒ その事象が生じる確率： $P = \lim_{n\to\infty} \frac{r}{n}$

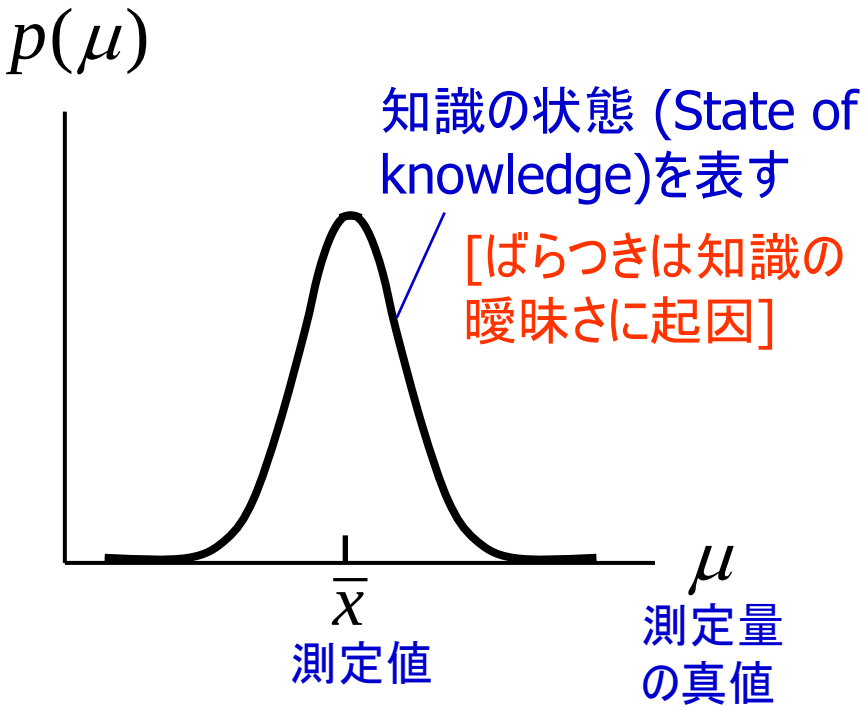
- 繰り返し観測可能な変数のみが確率変数になり得る
- 物理定数自体や過去の天気を確率変数と考えることはできない

確率概念の違い

頻度主義

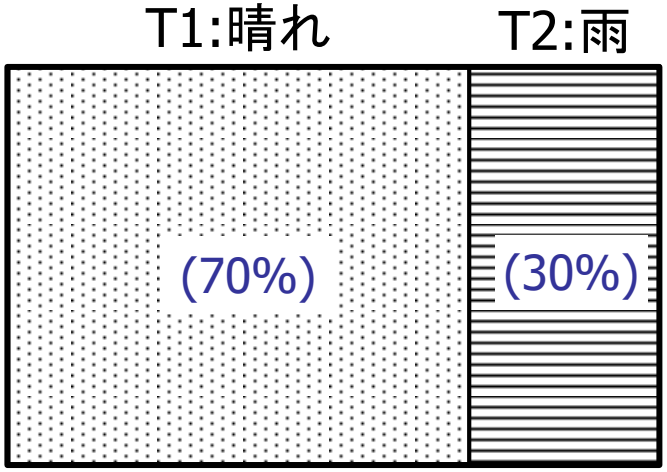


ベイズ統計



ベイズの定理

例）あるサッカーチームXの勝敗とスタジアムの天気



スタジアム周辺の過去の天気の割合

Q: 過去のある日(Day Z)が「雨」だった確率は？　　⇒ 30%

ベイズの定理(続き)

天気別の勝敗率がわかると･･･

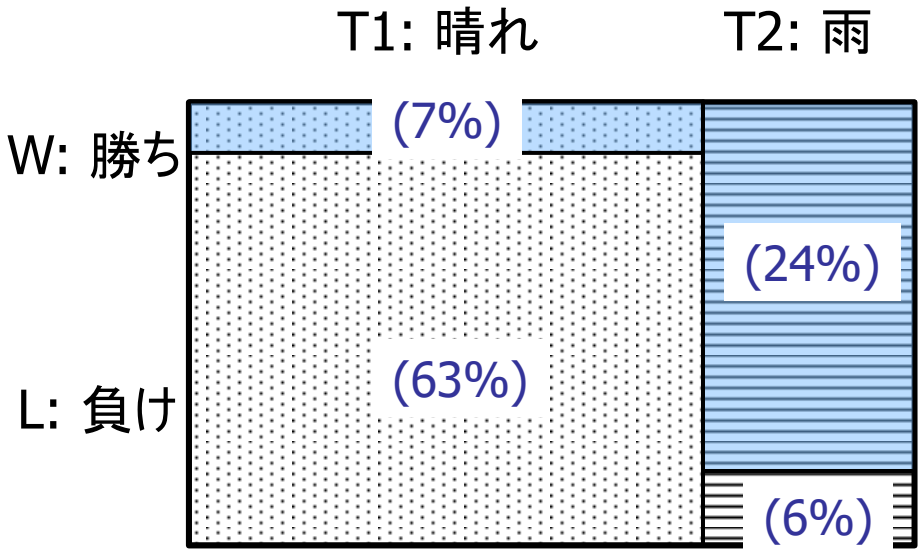
晴れのとき･･･勝10%, 負90%
雨のとき　･･･勝80%, 負20%

(チーム X は雨に強い)

⇩

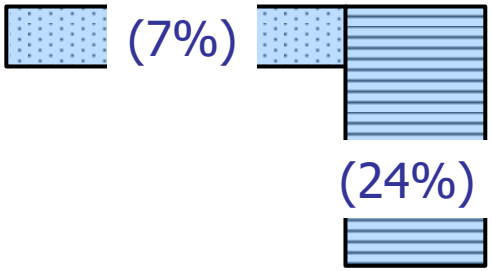
天気と勝敗の同時確率がわかる

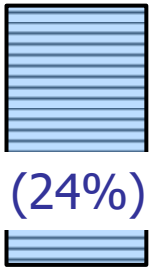
晴＆勝･･･　7%　晴＆負･･･63%
雨＆勝･･･24%　雨＆負 ･･･ 6%



ベイズの定理(続き)

Q: Day Zは「勝ち」であった。その日が「雨」だった確率は？



の内の 

の割合

$$= \frac{0.24}{0.7 + 0.24}$$
$$= 0.774 \ (\cong 77\%)$$

式で書くと・・・

雨という条件下での勝の確率

$$\Pr(\text{雨} \mid \text{勝}) = \frac{\Pr(\text{勝} \mid \text{雨}) \Pr(\text{雨})}{\Pr(\text{勝} \mid \text{雨}) \Pr(\text{雨}) + \Pr(\text{勝} \mid \text{晴}) \Pr(\text{晴})}$$

勝という条件下での雨の確率

・・・ベイズの定理

❑ 因果関係（天気⇒勝敗）を逆転した推論（勝敗⇒天気）が可能

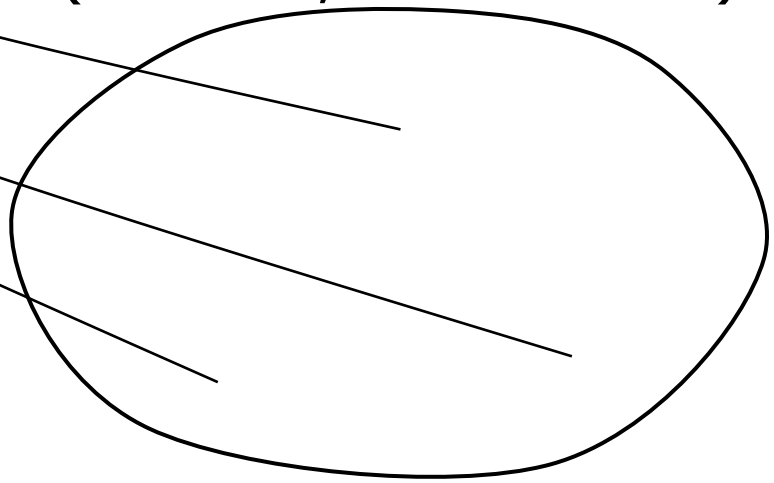
タイプA評価へのベイズ統計の利用

ある物質の濃度 C を繰り返し測定し、次のデータを得た。

標本　$\mathbf{x} = (x_1, x_2, \ldots, x_n)$

（標本平均: $\bar{x}$
標本標準偏差: s）

母集団
(母平均C, 母標準偏差σ)



$\mathbf{x}$ が決まった後の C の確率分布
（幾つかの仮定の下に）

$$p(C \mid \mathbf{x}) = T_{n-1}\left(\frac{C - \bar{x}}{s/\sqrt{n}}\right)$$

・・・ 自由度 n-1 の t分布

(scaled and shifted t-distribution)

濃度Cの標準不確かさ (タイプA評価)

ベイズ統計(GUM改訂案)

$$u_{\mathrm{Bayes}}(C) \;=\; p(C \mid \mathbf{x})\text{の標準偏差} \;=\sqrt{\frac{n-1}{n-3}}\,\frac{s}{\sqrt{n}}$$

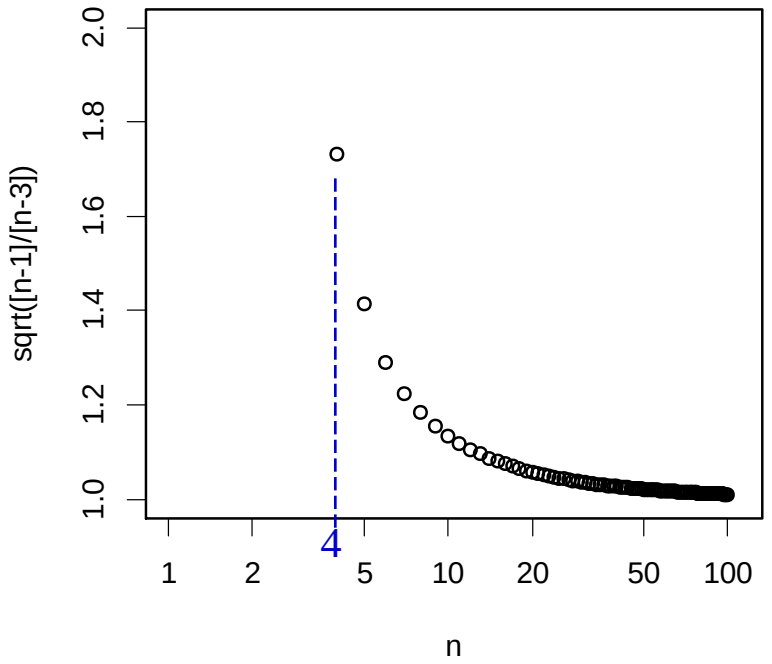
この分が異なる（必ず大きくなる）

頻度主義 (現行GUM)

$$u(C) = u(\bar{x}) = \frac{s}{\sqrt{n}}$$

因子 $\sqrt{(n-1)/(n-3)}$ の効果

n	$\sqrt{[(n-1)/(n-3)]}$
4	1.732
5	1.414
6	1.291
7	1.225
8	1.183
9	1.155
10	1.134
15	1.080
20	1.057
25	1.044
30	1.036
40	1.027
50	1.021
100	1.010
500	1.002
1000	1.001



- 繰り返し数 n が小さいときに、標準不確かさを大きくする
- $n = 3$ 以下では計算できない （自由度2以下の t 分布の性質）
  （その際は、プールされた実験標準偏差などを用いる）

タイプB評価

- 実務上ほとんどの場合に、現行GUMからの変更はない
  - 自由度 ∞ との想定ができるとき（あるいは利用情報の信頼性についての言及がないとき）: $u \rightarrow u$ （同じ）

- ただし u の信頼性(の欠如)についての情報がある場合、しかるべき修正を加える
  - u に自由度 ν が付与されているとき: $u \rightarrow \sqrt{\nu/(\nu-2)} \times u$
  - 利用情報の信頼性が見積もられているとき: ･･･事例集に例示

改訂案における包含区間の決定(1)

以下1)～3)のいずれかを用いる:

1) 分布形に依存しない包含係数 (少なくとも p の包含確率)
   - 任意の分布に対し: $k = 1/\sqrt{1-p}$ $\left[= 4.47 \quad \text{for} \quad p = 0.95 \right]$
   - 対称な単峰分布のとき: $k = 2/(3\sqrt{1-p})$ $\left[= 2.98 \ \text{for} \ p = 0.95 \right]$

   [根拠: Chebyshevの不等式、Gaussの不等式]

2) 確率分布から包含区間を決定
   (a) 測定量に対する知識の状態が正規分布で表される場合、正規分布表を利用 $k = 1.96$ for $p = 0.95$
   (b) 不明のときは、Monte Carlo法による分布の伝播(GUM 補完文書1[JCGM 101(2008)]が使える

改訂案における包含区間の決定(2)

3) 解析的に計算

1つの入力量だけが支配的な場合

確率変数　～定数

$$Y = f(X_1, x_2, \ldots, x_N)$$

$$p(y) = \left| \frac{1}{\partial f / \partial x_1} \right| p(x_1)$$

・・・ X_1 の確率分布を Y の確率分布に変換できる

「自由度」の扱い

頻度主義

- 自由度　＝　実験分散の（実験毎の）ばらつきを表す尺度

ベイズ統計

- 自由度の概念は現れない
- 手持ちの情報はすべて確率分布として表現され、情報の曖昧さは、確率分布自体に盛り込まれる

- ただし、Aタイプ評価の対象量の周辺事後分布として、自由度(n-1)の t 分布が自然に現れることがある
- この際の (n-1)は単に t分布のパラメータとしての意味しかもたず、情報の曖昧さの指標ではない

「ベイズらしさ」のレベル

確率概念

頻度主義統計
- タイプA評価だけのGUM

… 頻度基準

現行GUM
- 全体の論理は頻度主義
- タイプB評価に主観確率を導入

… 折衷

GUM改訂案
- タイプA評価をベイズ化
- 確率分布 p(y)の計算は必須でない

GUM Supplement1
- 骨格はベイズ統計
- 確率分布の計算はベイズ定理の代わりに「分布の伝播則」

ベイズ統計
- Full Bayes = 観測方程式 ＋ ベイズ定理＋周辺化

… 主観確率

GUM改訂案の実務上の主な影響

- タイプA評価(繰り返し数 n)における標準不確かさ

$$\frac{s}{\sqrt{n}} \text{ (現行GUM)} \quad \Rightarrow \quad \sqrt{\frac{n-1}{n-3}}\frac{s}{\sqrt{n}} \text{ (改訂案)}$$

- （有効）自由度の概念は消滅。Welch-Satterthwaiteの式は無用に

- 不確かさ伝播則は、継続して使用

- 包含係数 $k = 2$ が使いやすくなる

(繰り返し数n が小さくとも，$U = k\,u(y)$ において，$u(y)$自体が大きくなるため，単純に $k = 2$ としても $\pm U$ の信頼の水準が95 %より小さくなりにくい)

GUM改訂案に対する代表的な意見

C. Michotte資料 (BIPM Workshop)の抜粋等

否定的意見

- 改訂のメリットが明瞭でない
- 改訂に伴う社会的コストが大きすぎる
- 内容がむつかしい
- 最初に仮定する分布（先験的分布）の妥当性を検証しないまま，そのあとの論理だけを精緻化しようとしている
- 「不確かさは，測定誤差の大きさの尺度」とのもっともな期待から離反している

肯定的意見

- タイプAとタイプBの間の概念上の整合性が向上
- 改訂案のほうが理解しやすい
- 妥当性が低い有効自由度の計算が不要になる

・・・ 否定的意見が顕著に優勢であった

まとめ

- ベイズ統計を取り入れた、GUMの改訂作業が進められている
- 最初の改訂案(2014/12)についての関係各機関の意見は，改訂に否定的であった
- JCGMは今後，
  - 改訂のメリット・根拠についての説明・広報活動の強化
  - 現行GUMと改訂GUMを一定の遷移期間の間，共存させる仕組みの検討

  などを進める可能性がある
- 今後の予測は困難であるが，1~2年内のGUM改訂の可能性は低くなったと考えられる

GUM改訂に関する公開文献

論文

1) W. Bich, “How to revise the GUM?” *Accred. Qual. Assur.* 13 (2008) 271-275.
2) W. Bich, *et al.*, “Revision of the Guide to the Expression of Uncertainty in Measurement,” *Metrologia* 49 (2012) 702–705.
3) H. Imai, "Expanding needs for metrological traceability and measurement uncertainty," *Measurement* 46 (2013) 2942–2945.
4) 今井秀孝, 「JCGMの最新動向：GUM及びVIMの現状と将来」, 計測標準と計量管理, 65, No.2(2015) 52-59.

講演資料

1) BIPM Workshop on Measurement Uncertainty (June 15-16, 2015) [www.bipm.org/en/conference-centre/bipm-workshops/measurement-uncertainty/]

謝辞

今井秀孝様　（産業技術総合研究所、
製品評価技術基盤機構）
小池昌義様　（産業技術総合研究所）
田中秀幸様　（産業技術総合研究所）
城野克広様　（産業技術総合研究所）