包含係数の決定及び有効自由度の評価に関する資料

独立行政法人製品評価技術基盤機構
認定センター　計量認定課

1.　目的

この資料は、「JCSS 登録の一般要求事項（認定－部門－JCRP21）」5.2.1 項で定める校正証明書に記載する測定の不確かさの表現について、信頼の水準約 95 %に対応する区間を与える包含係数の決定及び／又は有効自由度の評価の考え方について検討し、JCSS 登録事業者における信頼の水準約 95 %に対応する拡張不確かさの決定に関する指針を与えることを目的としています。

2.　適用範囲

この資料は、JCSS 登録事業者が、測定の不確かさを拡張不確かさにより校正証明書に表記する次の(1)の場合に適用することができます。

【この資料で取り扱うケース】

(1)包含係数「2」が信頼の水準約 95 %に対応する区間を与えるとみなすことができ、$k = 2$ を採用してもよい場合。これには 2 つのケース即ち、

①Welch-Satterthwaite の式で計算しなくても有効自由度 ν_{eff} が十分大きいことが明らかな場合と、

②同式で計算した結果、有効自由度 ν_{eff} が十分大きいことが確認できた場合とがある。

この資料では、次の(2)の場合は当面取り扱いませんが、分野ごとに定める技術的要求事項適用指針又は不確かさの見積もりに関するガイドに特段の取決めがある場合には、それを適用することができます。

【この資料で現在取り扱わないケース】

(2) 出力量の確率分布が正規分布であるとみなせない場合、不確かさの伝播則に含まれる線形近似が適切でない場合、入力量に相関がある場合（注）、モンテカルロ法などの数値的方法による不確かさの合成法を採用する場合等、GUM 付属書 G に従って有効自由度を計算することが困難な場合。

なお、この資料の適用は、JCSS 登録事業者が GUM 付属書 G に従って厳密に有効自由度を計算し、信頼の水準約 95 %に対応する区間の包含係数を決定することを妨げるものではありません。

（注）相関のある入力量についてタイプ A 評価を用いている場合は、通常 GUM 4.1.4 注記の方法を用いることによって相関を考慮する必要がなくなります。

3.　引用規格及び関連文書

3.1　引用規格

ISO/IEC Guide 98-3:2008 Guide to the Expression of Uncertainty in Measurement（GUM）（測定における不確かさの表現のガイド）

3.2　関連文書

JCSS 登録の一般要求事項（認定－部門－JCRP21）
校正方法と不確かさに関する表現（認定－部門－JCG200）

4.　一般原則

JCSS 校正証明書で報告する測定結果の不確かさは、GUM 7.2.3 に従って報告することを原則とします。具体的には次のとおりです。

> 不確かさの尺度が拡張不確かさ $U = ku_{\mathrm{c}}(y)$ である場合、次の記述が望ましい。
> a) 測定量 Y がどのように定義されたかを完全に記述すること。
> b) 測定の結果を $Y = y \pm U$ と記述し、y と U に単位を付けること。
> c) 適切ならば、相対拡張不確かさ $U/|y|$、$|y| \neq 0$ を含めること。
> d) U を求めるために用いた k の値を与えること（又は結果の使用者の便宜のため、k と $u_{\mathrm{c}}(y)$ の両方を与えること）。
> e) 区間 $y \pm U$ に関連するおおよその信頼の水準を与え、またそれがどのように決められたかを述べること。

5.　包含係数決定に係る一般指針

5.1　包含係数 k=2 の採用

GUM 付属書 G.6.6 によれば、広範な分野における実際の多くの測定に対して、次の条件下では「$k = 2$ を採用し、$U = 2u_{\mathrm{c}}(y)$ が約 95 %の信頼の水準をもつ区間を定めるとみなすこと。」ができるとされています。多くの校正作業において、この条件は満たすと考えて良く、有効自由度 ν_{eff} が 10 以上ある場合は、約 95 %の信頼の水準をもつ区間の包含係数 k として、2 を採用できます。

> －測定量 Y の推定値 y は、正規分布や一様分布のような振舞いのよく分かった確率分布によって記述できる十分な数の入力量 X_i の推定値 x_i から求められる。
> －タイプ A 又はタイプ B の評価のどちらかによって求められるこれらの推定値の標準不確かさ $u(x_i)$ は、測定結果 y の合成標準不確かさ $u_{\mathrm{c}}(y)$ に対し、同程度の大きさの寄与を持つ。
> －不確かさの伝播則に含まれる線形近似が適切である（GUM 5.1.2 及び GUM 付属書 E.3.1 項参照）。
> －$u_{\mathrm{c}}(y)$ の不確かさは、その有効自由度 ν_{eff} が十分に大きい、例えば 10 以上、であるためにかなり小さい。

5.2　各要因の標準不確かさの自由度 ν_i が全て 10 以上の場合の包含係数

評価された各要因の標準不確かさの自由度 ν_i が全て 10 以上の場合は、有効自由度 ν_{eff} が十分大きいことが明らかであり、包含係数 k として、2 を採用できます。

5.3　有効自由度が十分大きいことを計算して確認する場合の包含係数

5.2 項以外の場合には、GUM 付属書 G.6.4 に基づき有効自由度 ν_{eff} を Welch-Satterthwaite の式で計算します。計算の結果、有効自由度 ν_{eff} が 10 以上の場合は、包含係数 k として、2 を採用できます。

なお、合成標準不確かさ$u_\mathrm{c}(y)$と Welch-Satterthwaite の式との関係は、次のとおりです。

$$u_\mathrm{c}(y)=\sqrt{u_1^2(y)+u_2^2(y)+u_3^2(y)+\ldots+u_n^2(y)}$$

$$\nu_\mathrm{eff}=\frac{u_\mathrm{c}^4(y)}{\sum_{i=1}^{n}\frac{u_i^4(y)}{\nu_i}}=\frac{u_\mathrm{c}^4(y)}{\frac{u_1^4(y)}{\nu_1}+\frac{u_2^4(y)}{\nu_2}+\frac{u_3^4(y)}{\nu_3}+\ldots+\frac{u_n^4(y)}{\nu_n}}$$

6.　実際的な考察

6.1　タイプ A 評価された標準不確かさの自由度

タイプ A 評価された標準不確かさの場合には、分散を算出するために用いた自由度を採用します。例えば、互いに独立なサンプル数 n が 10 個のデータであれば、自由度ν_iは $n-1$ から 9 となります。

6.2　タイプ B 評価された標準不確かさの自由度

ある標準不確かさ$u_\mathrm{i}(y)$がタイプ B 評価され、その標準不確かさの自由度ν_iに関する利用可能な情報が無い場合は、GUM 付属書 G.4.3 により$\nu_i\to\infty$（無限大）と仮定することは非現実的なことではありません。この場合$1/\nu_i\to 0$であり、5.3 項の下式のうち$u_i^4(y)/\nu_i\to 0$となります。

校正証明書から引用した拡張不確かさが、約 95 %の信頼の水準をもつ区間の包含係数として、2 以下の数が与えられている場合は、自由度ν_iは∞とすることができます。これは、包含係数 k が 2 の場合の自由度ν_iは、60 であり∞として扱うことに殆どの場合問題がないからです。なお、包含係数が 2 を超える場合は、t 分布表から有効自由度ν_effを求める必要があります。

（測定の不確かさの再評価において、上位の校正機関が発行した jcss 校正証明書又は JCSS 校正証明書に包含係数 k=2 の記載のみの場合等、有効自由度の記載がない場合は、当面、自由度ν_iは∞として扱って構いません。）

表１：有効自由度ν_effに対する約 95 %の信頼の水準をもつ区間の包含係数 k

ν_eff	1	2	3	4	5	6	7	8	9	10	20	50	∞
k	12.71	4.30	3.18	2.78	2.57	2.45	2.36	2.31	2.26	2.23	2.09	2.01	1.96

（包含係数 k は、t 分布表の$t_{95}(\nu_\mathrm{eff})$の値である）

6.3　Welch-Satterthwaite の式の計算例

計算例として、ケース 1 からケース 4 の事例を紹介します。

この事例の前提条件として、標準不確かさと自由度 ν_i、有効自由度 ν_{eff} の関係を容易に理解していただくため、

①不確かさの要因の値、標準不確かさを相対値（%）で表記し、感度係数もすべて 1 としています。実際の不確かさの評価では、不確かさの値の単位が異なる場合は、それぞれの感度係数により単位を統一して標準不確かさを算出してください。

②すべてのケースで不確かさの要因を 3 つとし、標準不確かさも、0.056 %、0.078 %、0.029 %の 3 つの組み合わせで、合成標準不確かさはすべて 0.100 %としています。「不確かさの要因　●●●」は、矩形分布として評価する標準器の表示分解能等、いくつかの事例が想定されますが、今回は便宜上、●●●としています。

③ケース 1、ケース 3 では 3 回の繰り返し測定の確率分布を正規分布と仮定しています。

④不確かさバジェット表は事例です。このスタイルに限定されるわけではありません。

ケース 1：タイプ B 評価された標準不確かさが支配的で、タイプ A 評価された標準不確かさが合成標準不確かさの半分程度以下の場合

例えば、$n=3$回の独立な測定の繰り返し性の相対標準不確かさ$u_1(x)$のみがタイプ A 評価され（自由度$\nu_1=3-1=2$）、その他の相対標準不確かさ$u_2(x)$及び$u_3(x)$はいずれもタイプ B 評価されており、相対標準不確かさ$u_2(x)$が支配的な場合を考えてみます。相対合成標準不確かさ$u_c(y)$の不確かさバジェットは次のようになります。

（ケース 1）

記号	不確かさの要因	タイプ	値（%）	確率分布	除数	標準不確かさ（%）	感度係数	標準不確かさ（%）	自由度 ν_i
$u_1(x)$	測定の繰り返し性（実験標準偏差による）	A	0.097	正規	$\sqrt{3}$	0.056	1	0.056	2
$u_2(x)$	標準器の校正（校正証明書による）	B	0.156	正規	2	0.078	1	0.078	∞
$u_3(x)$	●●●	B	0.05	矩形	$\sqrt{3}$	0.029	1	0.029	∞
$u_c(y)$	合成標準不確かさ							0.100	20

この場合においては、有効自由度$\nu_{\mathrm{eff}}=20$となり、包含係数$k=2$が採用できます。

$$\nu_{\mathrm{eff}}=\frac{0.100^4}{\left(\frac{0.056^4}{2}+\frac{0.078^4}{\infty}+\frac{0.029^4}{\infty}\right)}=\frac{0.100^4}{\left(\frac{0.056^4}{2}+0+0\right)}=20$$

［ポイント］

このケースは、$n=3$回の独立な測定の繰り返しの実験標準偏差が小さく、合成標準不確かさ 0.100 %に対して、その半分程度の 0.056 %の場合です。この場合は、$n=3$回の少ない測定回数であっても有効自由度$\nu_{\mathrm{eff}}=20$となり、包含係数$k=2$が採用できます。

ケース 2：<u>タイプ A 評価された標準不確かさが支配的であるが、その自由度が比較的大きい場合</u>

例えば、$n=8$ 回の独立な測定の繰り返し性の相対標準不確かさ $u_1(x)$ がタイプ A 評価され（自由度 $\nu_1=8-1=7$）、かつ、支配的であり、その他の相対標準不確かさ $u_2(x)$ 及び $u_3(x)$ はいずれもタイプ B 評価されている場合を考えてみます。相対合成標準不確かさ $u_\mathrm{c}(y)$ の不確かさバジェットは次のようになります。

（ケース 2）

記号	不確かさの要因	タイプ	値（%）	確率分布	除数	標準不確かさ（%）	感度係数	標準不確かさ（%）	自由度 ν_i
$u_1(x)$	測定の繰り返し性（実験標準偏差による）	A	0.221	正規	$\sqrt{8}$	0.078	1	0.078	7
$u_2(x)$	標準器の校正（校正証明書による）	B	0.112	正規	2	0.056	1	0.056	∞
$u_3(x)$	●●●	B	0.05	矩形	$\sqrt{3}$	0.029	1	0.029	∞
$u_\mathrm{c}(y)$	合成標準不確かさ							0.100	19

この場合においては、<u>有効自由度 $\nu_\mathrm{eff}=19$ となり、包含係数 $k=2$ が採用できます。</u>

$$\nu_\mathrm{eff}=\frac{0.100^4}{\left(\frac{0.078^4}{7}+\frac{0.056^4}{\infty}+\frac{0.029^4}{\infty}\right)}=\frac{0.100^4}{\left(\frac{0.078^4}{7}+0+0\right)}=19$$

［ポイント］

このケースは、測定の繰り返しの標準不確かさが大きく、合成標準不確かさが測定の繰り返しの標準不確かさの影響を大きく受ける場合です。この場合は、測定回数を増やすことで有効自由度 ν_eff が大きくなり（有効自由度 ν_eff が 10 以上）、包含係数 $k=2$ が採用できます。

ケース 3：タイプ A 評価された標準不確かさが支配的であり、その自由度が小さい場合

例えば、$n=3$回の独立な測定の繰り返し性の相対標準不確かさ$u_1(x)$がタイプ A 評価され（自由度$\nu_1=3-1=2$）、かつ、支配的であり、その他の相対標準不確かさ$u_2(x)$及び$u_3(x)$はいずれもタイプ B 評価されている場合を考えてみます。相対合成標準不確かさ$u_c(y)$の不確かさバジェットは次のようになります。

（ケース 3）

記号	不確かさの要因	タイプ	値（%）	確率分布	除数	標準不確かさ（%）	感度係数	標準不確かさ（%）	自由度 ν_i
$u_1(x)$	測定の繰り返し性（実験標準偏差による）	A	0.135	正規	$\sqrt{3}$	0.078	1	0.078	2
$u_2(x)$	標準器の校正（校正証明書による）	B	0.112	正規	2	0.056	1	0.056	∞
$u_3(x)$	●●●	B	0.05	矩形	$\sqrt{3}$	0.029	1	0.029	∞
$u_c(y)$	合成標準不確かさ							0.100	5.4

この場合においては、有効自由度$\nu_{\text{eff}}=5.4$となり包含係数$k=2$を採用すると過小評価となります。t分布の$t_{95}(5)$の値は「2.57」となります。

$$\nu_{\text{eff}}=\frac{0.100^4}{\left(\dfrac{0.078^4}{2}+\dfrac{0.056^4}{\infty}+\dfrac{0.029^4}{\infty}\right)}=\frac{0.100^4}{\left(\dfrac{0.078^4}{2}+0+0\right)}=5.4$$

［ポイント］

このケースは、ケース２と同様に測定の繰り返しの標準不確かさが大きく、合成標準不確かさが測定の繰り返しの標準不確かさの影響を大きく受ける場合です。しかし、この例のように測定回数を増やすことができない場合は、有効自由度ν_{eff}が小さくなり、包含係数$k=2$が採用できないため、t分布表から包含係数k*を求める必要があります。

（*：2011 年 6 月 20 日　「有効自由度ν_{eff}」を「包含係数 k」に訂正しています。）

ケース 4：過去にプールされた実験標準偏差を利用し、タイプ A 評価をする場合

例えば、ケース 3 の場合、統計的管理状態に保たれている、はっきりと素性が知られた測定に対しては、測定を特徴づける合成又はプールされた分散の推定値 $s_{\mathrm{p}}{}^2$（多数回（以下の例では 10 回）の測定に基づきプールされた実験標準偏差 s_{p}）を利用できることがあります。このような場合、測定値が独立な n 個の観測値から決められるときには、観測値の相加平均 $\bar{q}$ の実験分散は $s^2(\bar{q})/n$ よりも、$s_p{}^2/n$ によってより良く推定され、標準不確かさは $u = s_p/\sqrt{n}$ となります。

校正品目とプールデータの一致性が確認された場合、プールされた実験標準偏差 s_{p} を利用し、$n=3$ 回の独立な測定の繰り返し性の相対標準不確かさ $u_1(x)$ は $u_1(x)=s_p/\sqrt{n}$ で求めることができ、自由度 $\nu_1 = 10-1=9$ となります。この場合の相対合成標準不確かさ $u_{\mathrm{c}}(y)$ の不確かさバジェットは次のようになります。(GUM 4.2.4 参照)

（ケース 4）

記号	不確かさの要因	タイプ	値（%）	確率分布	除数	標準不確かさ（%）	感度係数	標準不確かさ（%）	自由度 ν_i
$u_1(x)$	測定の繰り返し性 (プールされた実験標準偏差による)	A	0.135	正規	$\sqrt{3}$	0.078	1	0.078	9
$u_2(x)$	標準器の校正 (校正証明書による)	B	0.112	正規	2	0.056	1	0.056	∞
$u_3(x)$	●●●	B	0.05	矩形	$\sqrt{3}$	0.029	1	0.029	∞
$u_{\mathrm{c}}(y)$	合成標準不確かさ							0.100	24.3

この場合においては、有効自由度 $\nu_{\mathrm{eff}} = 24.3$ となり、$k=2$ が採用できます。

$$\nu_{\mathrm{eff}} = \frac{0.100^4}{\left(\dfrac{0.078^4}{9}+\dfrac{0.056^4}{\infty}+\dfrac{0.029^4}{\infty}\right)} = \frac{0.100^4}{\left(\dfrac{0.078^4}{9}+0+0\right)} = 24.3$$

［ポイント］

このケースは、ケース３（測定回数を増やすことができない場合）であっても、過去に同じ校正品目（例えば、同一型式の計測器）でプールされた実験標準偏差（0.135 %）があり、その値と $n=3$ 回の独立な測定の繰り返しの実験標準偏差がほぼ同程度と判断できる場合は、プールされた実験標準偏差を利用することができる事例です。

$n=3$ 回の独立な測定の繰り返しですが、プールされた実験標準偏差を求めた際の測定回数 10 回から自由度 $\nu_1 = 10-1=9$ として計算することができます。その結果、$n=3$ 回の少ない測定回数であっても有効自由度 $\nu_{\mathrm{eff}} = 24.3$ となり、包含係数 $k=2$ が採用できます。

このケースでは「校正品目とプールデータの一致性の確認」を同一型式の計測器としましたが、どこまで許容できるのか今後 JCSS 分野別技術分科会で検討する予定です。

7.　その他

タイプ A 評価された標準不確かさから合成標準不確かさを求める場合、測定回数があまり少ないと「不確かさの不確かさ」が大きくなってしまいます（GUM 表 E.1 参照）。

これを回避するために、幾つかの方法がありますが、最終的にどのような測定方法を採用し、測定値のばらつきを示すバロメータである不確かさを評価するかは、各校正事業者の判断にゆだねられます。以下に、幾つかの事例について、メリット及びデメリット若しくはポイントを示します。

(1) 十分な測定回数（例：10 回以上）による測定を行う

メリット：	測定の繰り返し回数が多くなるため、不確かさの不確かさが小さくなる方向に寄与し、合成標準不確かさの有効自由度 ν_{eff} も大きくなる（$k = 2$ が採用しやすくなる）。
デメリット：	校正事業者におけるコストとの収支バランスを考えると、現実的な校正作業とは言えない場合がある。

(2) 事前にプールデータをとっておき、個々の校正では少ない測定回数を採用する（6.3　ケース 4 参照）

メリット：	①顧客の混乱や、校正事業者の負担増は想定されない。 ②少ないサンプル数のデータによる不確かさ評価より、良い推定ができる。 ③同一の校正品においては、同じ不確かさを用いることができる。 ④製品管理データが活用できる（利用可能な場合）。
ポイント：	①その測定量の測定開始当初は、プールデータをどの程度の測定回数から評価すれば良いか、分野ごと、計測器ごとの相場観を元に、個々の校正での（少ない）測定回数との合理的な組み合わせを含めて検討が必要となる。 ②依頼される校正品目とプールデータの一致性があることの確認が必要となる。

(3) 測定の都度、有効自由度 ν_{eff} を評価する

メリット：	①k=2 を用いるよりも、適切な信頼の水準の区間を与えることができる。 ②k=2 を用いるよりも、拡張不確かさの過小評価のリスクを低減できる。
デメリット：	①校正事業者における不適合業務を誘発しかねず、予防処置のための作業負担（例：ソフトウェアの修正・保護）が増す。また、顧客の混乱も懸念される。 ②「不確かさの不確かさ」が非常に大きいため、真の不確かさの値と非常に異なる値が算出されている可能性がある。

以　上