既知のかたよりを補正しない場合の不確かさ評価に関する一考察* —ブロックゲージ比較測定法を例として—

尾 藤 洋 一** 榎 原 研 正***

Consideration for Uncertainty Evaluation Method with Uncorrected Deflection

Youichi BITOU and Kensei EHARA

We consider an uncertainty evaluation method with uncorrected deflection. In the real uncertainty evaluation, there are a lot of cases which do not correct the deflection and this deflection is simply added to the total uncertainty in the conventional evaluation method. In contrast, we clarify the uncertainty evaluation method with uncorrected deflection along a way of thinking of GUM. The conventional uncertainty evaluation method is overestimated compared with that along a way of thinking of GUM.

Key words: uncertainty, deflection, GUM

1. 緒 言

測定結果の信頼性の表現方法として、近年、「不確かさ」(Uncertainty)が広く用いられるようになってきている。例えば、わが国におけるJCSS(Japan Calibration Service System)制度等、測定結果のトレーサビリティを国際的に確保するためのしくみの中では、「精度」や「誤差」ではなく、「不確かさ」が用いられている。「不確かさ」の定義およびその評価方法は、「計測における不確かさの表現のガイド」(GUM)¹⁾²⁾に規定されている。

GUMに従って測定結果の不確かさを求める際には、様々な不確かさの要因毎に、標準偏差に相当する大きさの不確かさ、すなわち「標準不確かさ」を見積もり、それらを不確かさの伝播則によって合成することにより「合成標準不確かさ」を求める。個々の要因毎の不確かさは、実験による測定データの変動から統計的に求めるAタイプ評価、もしくは校正証明書や既存の文献などの外部情報に基づいて推定するBタイプ評価のいずれかにより見積もられる。しかし、現実には、全ての不確かさ成分をGUMに則った手順で容易に評価できるわけではない。その場合、個々の事例ごとに、GUMの考え方に出来るだけ沿う形で不確かさの評価法を検討する必要がある。しかし、GUMの考え方と必ずしも一致しない不確かさ評価が慣例化してしまっている例がある。幾何学量(長さ)計測における、かたよりを補正しない場合の不確かさ評価もこの一例である。

かたよりを補正しない不確かさ評価法とは、例えば何らかの参照標準を用いて測定器の校正を行う場合に、参照標準に対する読み値と参照標準に付与された値(標準値)に差があることがわかっていても、それを補正せず、不確かさの一分成分として考慮する方法を意味する。GUMでは、このような既知のかたよりのある場合は、原則として補正を行うことを薦めている。しかし、

Fig.1 Gauge block comparator

- (1) かたよりの大きさが推定可能であっても、その値自体の不確かさが大きいため、補正することの技術的妥当性が不明瞭(後述する平均温度測定の場合、あるいはいわゆるゼロガスによる校正やコサイン誤差など、かたより自体が何らかの誤差要因により生じる場合)
- (2) コストや労力などの点で補正しない方が有利、あるいは補正が技術的に困難(補正機能を内蔵しない測定器の場合など)

などの理由から、かたよりを補正せずに不確かさに含める場合がある。特に幾何学量計測において、このような例は少なくない。かたよりを補正しないこと自体は実用上許容されるべきであるが、その際の不確かさ評価の手順がGUMには明瞭に規定されていないために、このような場合の不確かさ評価方法には、しばしば混乱が生じている。さらに、GUMの考え方とは必ずしも整合しない評価方法が常態化している場合もある。

本報告では、ブロックゲージの比較測定を例に、かたよりを補正しない場合の、GUMの考え方に沿った不確かさ評価方法を示すとともに、これと従来からしばしば利用され

* 原稿受付 平成 19 年 10 月 4 日

** 正 会 員 産業技術総合研究所(茨城県つくば市梅園 1-1-1)

*** 産業技術総合研究所

てきた評価方法（従来法と呼ぶ）との違いを明らかにする。そして、従来法を適用する場合の問題点、留意点について述べる。

2. ブロックゲージの比較測定

ブロックゲージ比較測定器の構成図を図 1 に示す。被校正用ブロックゲージの長さ L は、標準ブロックゲージ（校正証明書に記載された長さを L_s とする）との比較測定の結果 d を用いて、近似的に

$$L = L_s + d - L_s(\delta\alpha \cdot \theta + \alpha_s \cdot \delta\theta) \quad (1)$$

と与えられる³⁾。ここで、 $\delta\alpha$ は 2 本のブロックゲージの熱膨張係数の差、 θ は被校正ブロックゲージ温度の基準温度 20℃からの偏差、 α_s は標準ブロックゲージの熱膨張係数、 $\delta\theta$ は 2 本のブロックゲージの温度差である。通常の比較測定においては、同材質のブロックゲージを同一環境で測定するので、 $\delta\alpha = 0$ 、 $\delta\theta = 0$ として、 $L = L_s + d$ の関係式から被校正ブロックゲージの長さが求められる。一方、 L の不確かさ $u_c(L)$ の評価では、 $\delta\alpha$ 、 $\delta\theta$ に付随する不確かさも考慮して、式(1)から

$$u_c^2(L) = u^2(L_s) + u^2(d) + L_s^2 \alpha_s^2 u^2(\delta\theta) + L_s^2 \theta^2 u^2(\delta\alpha) + L_s^2 u^2(\alpha_s) u^2(\delta\theta) + L_s^2 u^2(\theta) u^2(\delta\alpha) \quad (2)$$

と与えられる⁴⁾。ここで、 $u(x)$ は入力量 x の標準不確かさを表す。式(2)右辺の不確かさ成分のうち、 $u(d)$ 、 $u(\delta\theta)$ 、 $u(\theta)$ に関しては、実際の不確かさ評価において、かたよりを補正しない状態での不確かさ評価が行われる場合が多い^{3)~5)}。

以下では、まず、比較測長に付随する不確かさ $u(d)$ の評価を例にとり、GUM の考え方に則った不確かさ評価法を提案するとともに、従来法との違いおよびその解釈について述べる。

3. 不確かさ評価の例 1：比較測長器

3.1 かたよりを補正する場合

ここではまず、比較測長器のかたよりを推定し、それを補正する場合の不確かさ評価について検討する。この場合の測定および不確かさ評価の手順として以下を想定する

(1) 測定対象と同等のブロックゲージ対を予め十分な回数（例えば 10 回以上）繰り返し測定し、実験標準偏差（不偏分散の平方根） s を求める。これは、プールされた実験分散にもとづく、標準不確かさの A タイプ評価である。

(2) 校正済みのブロックゲージ 2 個から構成した基準段差（段差の標準値 l_{ref} ）を準備する。 l_{ref} とその不確かさ u_{ref} は、いずれもブロックゲージの校正証明書に基づいて求める。基準段差を比較測長器で n_{ref} 回繰り返し測定し、その平均値を x_{ref} とする。 x_{ref} を用いて、比較測長器のかたより（系統誤差）の推定値 D_{bias} を次から決める。

$$D_{\text{bias}} = x_{\text{ref}} - l_{\text{ref}} \quad (3)$$

(3) 実際の測定対象となる段差試料を比較測長器で n 回繰り返し測定し（通常は $n = 2$ 程度）、その平均値を x とする。 x と D_{bias} から、かたよりを補正した後の測定結果 d_c を次式で求める。

Fig. 2 Probability density distributions of ε_c and ε_{nc} assumed in the conventional evaluation of uncertainty.

$$d_c = x - x_{\text{ref}} + l_{\text{ref}} \quad (4)$$

式(4)において、 x と x_{ref} には、 D_{bias} で推定されるかたより成分が共通に含まれるが、これらは差 $x - x_{\text{ref}}$ をとることにより相殺されるため、 x と x_{ref} の不確かさとしては、繰り返し測定のばらつきに付随する不確かさのみを考慮すればよいことに注意する。 d_c の不確かさ $u(d_c)$ は、式(4)に不確かさの伝播則を適用することにより、次のように求まる（付録 A 参照）。

$$u^2(d_c) = \left(\frac{s^2}{n_{\text{ref}}} + u_{\text{ref}}^2 \right) + \frac{s^2}{n} \quad (5)$$

ここで、右辺第 1 項は補正に伴う不確かさを、第 2 項は測定のばらつきに伴う不確かさを表している。式(5)に示される不確かさは、GUM に示されている評価法に完全にに基づいているものである。

3.2 かたよりを補正しない場合 ー従来法ー

かたよりの補正を行わない場合、段差の測定結果 d_{nc} としては、実際の測定における平均値 x をそのまま用いることになる。すなわち

$$d_{nc} = x \quad (6)$$

として測定対象の測定結果が与えられる。

ここで、かたよりを補正しない場合でも、測定結果の不確かさ評価のためには、前節の手順(1)~(3)がすべて必要であることに注意する。測定結果 d_c および d_{nc} に含まれる誤差をそれぞれ ε_c 、 ε_{nc} とするとき、 d_{nc} の不確かさ評価における従来法の一般的な考え方を図 2 に示す。 ε_c は期待値が 0、標準偏差が $u(d_c)$ の確率密度関数 $p(\varepsilon_c)$ に従うと想定することができ、また、 ε_{nc} の確率密度関数 $p(\varepsilon_{nc})$ は $p(\varepsilon_c)$ を D_{bias} だけ変位させたものと想定できる。従来法における d_{nc} の不確かさを $u_1(d_{nc})$ と表記すると、これは上のことから次のように評価される（付録 B 参照）。

$$\begin{aligned} u_1^2(d_{nc}) &= D_{\text{bias}}^2 + u^2(d_c) \\ &= D_{\text{bias}}^2 + \left(\frac{s^2}{n_{\text{ref}}} + u_{\text{ref}}^2 \right) + \frac{s^2}{n} \end{aligned} \quad (7)$$

実際の比較測長器の不確かさ評価では、通常、ばらつきの影響 s^2/n_{ref} が u_{ref}^2 と比較して無視できる程度に十分に大きな n_{ref} が選択される。この場合、式(7)は次のように近似される。

$$u_1^2(d_{nc}) = D_{\text{bias}}^2 + u_{\text{ref}}^2 + \frac{s^2}{n} \quad (8)$$

式(8)の導出は一見妥当に見えるが、不確かさの伝播則に従って厳密に導かれたものではない。 d_{nc} には影響を与えないはずの補正に伴う不確かさ(右辺第2項)が d_{nc} の不確かさの成分として合成されていることから、不確かさの評価方法として厳密性を欠くことは明かである。厳密には、次に示すような、データの構造モデルを用いた統計的検討が必要である。

3.3 かたよりを補正しない場合 -GUMに基づく方法-

ここでは、かたよりを補正しない場合の不確かさ評価として、GUMの考え方に従った、厳密な方法を検討する。ただし、GUMには既知のかたよりを補正しない場合の不確かさ評価の方法が具体的には規定されていないため、 x 、 x_{ref} および l_{ref} の統計的構造モデルについての検討から始める。

いま、測定対象試料および基準段差の真の段差値をそれぞれ λ 、 λ_{ref} とし、比較測長器が持つかたよりの真値を Δ_{bias} と表記する。これらはいずれも不可知量である。 x 、 x_{ref} および l_{ref} の構造モデルはそれぞれ次のように書ける。

$$x = \lambda + \Delta_{bias} + \varepsilon \quad (9)$$

$$x_{ref} = \lambda_{ref} + \Delta_{bias} + \varepsilon_{ref} \quad (10)$$

$$l_{ref} = \lambda_{ref} + \gamma \quad (11)$$

ここで、 ε および ε_{ref} はそれぞれ x および x_{ref} に含まれる誤差の内の偶然成分(ばらつき)を表し、 γ は l_{ref} がもつ誤差を表す。一般に、確率変数 y の期待値および分散をそれぞれ $E[y]$ 、 $V[y]$ のように表すことにすると、 ε および ε_{ref} について次が成立する。

$$E[\varepsilon] = E[\varepsilon_{ref}] = 0 \quad (12)$$

$$V[\varepsilon] = E[\varepsilon^2] = \frac{\sigma^2}{n} \quad (13)$$

$$V[\varepsilon_{ref}] = E[\varepsilon_{ref}^2] = \frac{\sigma^2}{n_{ref}} \quad (14)$$

ただし、 σ^2 は繰り返し測定にともなうばらつきの母分散で、 s と $E[s^2] = \sigma^2$ の関係にある。

γ は測定中に変動する量ではないが、GUMではこれを確率変数とみなし、何らかの先験的分布¹²⁾で記述できるものとする。この先験的分布の母分散を σ_{ref}^2 とし、確率変数 γ の期待値と分散を、混乱のない範囲で上と同じ記号 $E[\gamma]$ 、 $V[\gamma]$ で表すと、次が成立する。

$$E[\gamma] = 0 \quad (15)$$

$$V[\gamma] = \sigma_{ref}^2 \quad (16)$$

式(15)は、 l_{ref} が λ_{ref} の最良推定値¹²⁾と考えられることによる。

一般に、既知のかたよりが補正される場合、測定結果 y の不確かさの二乗 $u^2(y)$ は、統計的には、母分散 $E[(y-\mu)^2]$ の推定値として位置づけられる(GUM Annex E)。ここで $\mu = E[y]$ である。すなわち、測定結果が、推定すべき値からどれだけずれているかを、2乗平均として評価しようとするものが $u^2(y)$ である。今の場合、測定結果 d_{nc} が推定すべき値は λ であることから、 $u^2(d_{nc})$ は式(6)、(9)から、次の量の推定値と考えることができる。

$$E[(d_{nc} - \lambda)^2] = E[(\Delta_{bias} + \varepsilon)^2] = \Delta_{bias}^2 + \frac{\sigma^2}{n} \quad (17)$$

式(17)において、偶然成分 σ^2 は s^2 で推定できるから、 $u^2(d_{nc})$ の評価は系統的成分 Δ_{bias}^2 の推定に帰着する。

Δ_{bias}^2 を推定するため、式(10)、(11)から得られる D_{bias} の構造模型

$$D_{bias} = \Delta_{bias} + \varepsilon_{ref} - \gamma \quad (18)$$

を利用する。式(18)両辺の二乗の期待値をとり、項を一部入れ替えることにより

$$\Delta_{bias}^2 = E[D_{bias}^2] - \frac{\sigma^2}{n_{ref}} - \sigma_{ref}^2 \quad (19)$$

が得られる。ここで、 Δ_{bias} 、 ε_{ref} および γ の間に相関がないことを利用した。なお、前節で示したように、 σ_{ref}^2 は u_{ref}^2 で推定できる。また、式(18)から導かれる $\Delta_{bias} = D_{bias} - \varepsilon_{ref} + \gamma$ の両辺の二乗の期待値をとることにより、式(19)と矛盾する $\Delta_{bias}^2 = E[D_{bias}^2] + \sigma^2/n_{ref} + \sigma_{ref}^2$ が得られるように見えるが、 D_{bias} と ε_{ref} および γ の間には相関があるため(式(18)参照)、後者の式は成立しないことに注意する。式(19)から、 Δ_{bias}^2 は次により推定できる。

$$\widehat{\Delta_{bias}^2} = D_{bias}^2 - \frac{s^2}{n_{ref}} - u_{ref}^2 \quad (20)$$

ただし、 $\wedge$ は推定量であることを表す。式(20)は、 D_{bias}^2 は Δ_{bias}^2 の過大推定量であり、 Δ_{bias}^2 を不偏推定するには右辺第2および第3項が必要であることを意味する。しかし、比較測長器のかたよりが統計的に有意な大きさを持たない場合、実際の評価において、第2、3項のために式(20)右辺が負になることがあり得る。これは、例えば分散分析において、各要因ごとの不偏分散の期待値の表式を用いて母分散を推定するときに、結果が負の値になり得るのと同様の事情である。この場合には、式(20)に替えて

$$\widehat{\Delta_{bias}^2} = 0 \quad (21)$$

とすることを規約する。

従来法による不確かさ $u_I(d_{nc})$ と区別して、本節による不確かさを $u_{II}(d_{nc})$ とすると、式(17)、(20)、(21)より次を得る。

$$u_{II}^2(d_{nc}) = \begin{cases} D_{bias}^2 - s^2/n_{ref} - u_{ref}^2 + s^2/n, & \text{if } D_{bias}^2 - s^2/n_{ref} - u_{ref}^2 \geq 0 \\ s^2/n, & \text{otherwise} \end{cases} \quad (22)$$

ここで、前節と同様に、 s^2/n_{ref} が他の項に比べて無視できる場合、式(22)は次のように近似できる。

$$u_{II}^2(d_{nc}) = \begin{cases} D_{bias}^2 - u_{ref}^2 + s^2/n, & \text{if } D_{bias}^2 - u_{ref}^2 \geq 0 \\ s^2/n, & \text{otherwise} \end{cases} \quad (23)$$

3.4 不確かさ評価法の比較

ここでは、上に導いた不確かさ評価方法を従来法と比較し、それぞれの意味と問題点を検討する。特にかたよりの扱いに関心があるので、簡単のため、これ以降 σ^2 、 σ_{ref}^2 はそれぞれ s^2 、 u_{ref}^2 で十分正確に推定されていると仮定し($\sigma^2 = s^2$ 、 $\sigma_{ref}^2 = u_{ref}^2$)、さらに s^2/n_{ref} が他の項に比べて無視し得るもの($\varepsilon_{ref} = 0$)とする。この場合、式(8)、(23)はそれぞれ

$$u_I^2(d_{nc}) = D_{bias}^2 + \sigma_{ref}^2 + \sigma^2/n \quad (8')$$

$$u_{II}^2(d_{nc}) = \begin{cases} D_{bias}^2 - \sigma_{ref}^2 + \sigma^2/n, & \text{if } D_{bias}^2 - \sigma_{ref}^2 \geq 0 \\ \sigma^2/n, & \text{otherwise} \end{cases} \quad (23')$$

と表すことができる。

式(8'), (23')における $u_I^2(d_{nc})$, $u_{II}^2(d_{nc})$ を D_{bias}^2 の関数として図3に示す。 $u_I^2(d_{nc})$ は $u_{II}^2(d_{nc})$ よりも、 $\sigma_{ref}^2 \sim 2\sigma_{ref}^2$ だけ大きめの評価になっていることがわかる。この違いは、 D_{bias}^2 における σ_{ref}^2 の扱いの違いに起因している。 D_{bias}^2 は真の段差値の二乗 Δ_{bias}^2 より、平均して σ_{ref}^2 だけ大きい(式(19))ため、この差を修正したものが $u_{III}^2(d_{nc})$ である。一方、 $u_I^2(d_{nc})$ では、かたよりを補正する際に補正の不確かさとして生じる σ_{ref}^2 の項を、補正しないにもかかわらず不確かさの成分として合成しているため、不確かさとしては過大評価になっている。

既に述べたように、基準段差の誤差 γ は、手順(1)から(3)の間に変動する変数ではないが、GUM の考え方によると、 γ は何らかの先験的分布に従う確率変数とみなされる。 γ の分布は、 Δ_{bias} , l_{ref} および u_{ref} がすべて同一の比較測長器および基準段差を使用する多数の測定者、すなわち同一の測定能力を有するが γ の大きさはあらゆる可能な値をとる多数の測定者の分布とみなすことができる。 $u_{III}^2(d_{nc})$ はこのような多数の測定者について平均したときに、式(17)と一致する妥当な値を与える。すなわち、 $\Delta_{bias}^2 + \sigma^2/n$ の不偏推定量となっている。ただし、特定の測定者については、 $u_{III}^2(d_{nc})$ は過小評価にも過大評価にもなり得る。例えば、 γ の大きさ $|\gamma|$ が同じであっても、その符号が Δ_{bias} と同一か違うかに応じて、見かけのかたより D_{bias} が大きく違うことがあり得る。これは、それぞれの測定者が γ の値を個別に推定することができない以上、やむを得ない。

ここで、不確かさの二乗 $u^2(d_{nc})$ が $\Delta_{bias}^2 + \sigma^2/n$ の不偏推定量でないことを許容すれば、 $u_I(d_{nc})$ と $u_{II}(d_{nc})$ の中間をとった次の不確かさ評価式 $u_{III}(d_{nc})$ を考えることができる。

$$u_{III}^2(d_{nc}) = D_{bias}^2 + \frac{\sigma^2}{n} \quad (24)$$

$u_{III}(d_{nc})$ は、 $u_I(d_{nc})$ のような明かな過大評価を避けつつ、推定する不確かさが $\Delta_{bias}^2 + \sigma^2/n$ の不偏推定量から乖離することをできるだけ避けようとしたものと言える。 $u_I^2(d_{nc})$, $u_{II}^2(d_{nc})$ と同様に、 $u_{III}^2(d_{nc})$ を D_{bias}^2 の関数として図3に示す。図より、 $u_{III}(d_{nc})$ は $u_I(d_{nc})$ と $u_{II}(d_{nc})$ の中間に位置することが分かる。

式(24)で与えられる不確かさ $u_{III}^2(d_{nc})$ は、式(23)のように、 D_{bias} の値に依存した場合分けの必要がなく、表式が簡単という利点がある。さらに、 $u_{III}(d_{nc})$ は、GUM の考え方に忠実に沿った標準不確かさとは言えないものの、式(21)に示される様なかたよりの推定値が0となってしまう可能性は低いので、不確かさの過小評価が特に望ましくない状況では、現実的な選択肢となり得ると考えられる。なお、 σ^2 が s^2 により十分正確に推定されていると仮定できる場合、式(24)は次式で近似できる。

$$u_{III}^2(d_{nc}) = D_{bias}^2 + \sigma^2/n \quad (24')$$

Fig. 3 $u_I(d_{nc})$, $u_{II}(d_{nc})$ and $u_{III}(d_{nc})$ given in Eqs. (8'), (23') and (24') as functions of D_{bias}^2 .

3.5 不確かさが過小評価となる確率

前節で提示した3つの不確かさ評価法においては、式(8)には不確かさの過大評価となってしまう可能性が高い、式(23)には不確かさの過小評価となってしまう可能性があるといった問題点があると考えられる。実際に式(8), (23)および(24)がどの程度の頻度で過小評価(過大評価)を与える可能性があるかは、 $u^2(d_{nc})$ は $\Delta_{bias}^2 + \sigma^2/n$ と一致するときにも最も正確な推定となっているとの観点から、 Δ_{bias} を変数とし、 γ を標準偏差 σ_{ref} ($= u_{ref}$) の正規分布としたときの確率計算より見積もることが出来る。ここで簡単のため、近似式(8'), (23')および(24')を用いることとし、また、 σ_{ref} で規格化した変数

$$\tilde{\gamma} = \gamma / \sigma_{ref} \quad (25)$$

$$\tilde{\Delta}_{bias} = \Delta_{bias} / \sigma_{ref} \quad (26)$$

を計算に用いる。ここで、 $u_I(d_{nc})$, $u_{II}(d_{nc})$, $u_{III}(d_{nc})$ が過小評価になる確率をそれぞれ P_I , P_{II} , P_{III} とすると、式(17), (8') および式(18)より導かれる $D_{bias} = \Delta_{bias} - \gamma$ の関係式より、 $|\tilde{\Delta}_{bias}| > 1$ のとき、

$$\begin{aligned} P_I &= \Pr[(\Delta_{bias} - \gamma)^2 + \sigma_{ref}^2 + \sigma^2/n < \Delta_{bias}^2 + \sigma^2/n] \\ &= \Pr[\tilde{\Delta}_{bias} - \sqrt{\tilde{\Delta}_{bias}^2 - 1} < \tilde{\gamma} < \tilde{\Delta}_{bias} + \sqrt{\tilde{\Delta}_{bias}^2 - 1}] \\ &= \int_{\tilde{\Delta}_{bias} - \sqrt{\tilde{\Delta}_{bias}^2 - 1}}^{\tilde{\Delta}_{bias} + \sqrt{\tilde{\Delta}_{bias}^2 - 1}} \phi(\tilde{\gamma}) d\tilde{\gamma} \end{aligned} \quad (27)$$

と与えられる。ここで、 $\Pr$ は確率を表し、 $\phi(\tilde{\gamma})$ は規準正規分布

$$\phi(\tilde{\gamma}) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\tilde{\gamma}^2}{2}\right) \quad (28)$$

である。また、 $|\tilde{\Delta}_{bias}| < 1$ のときには、式(8')より、明らかに $u_I^2(d_{nc}) > \Delta_{bias}^2 + \sigma^2/n$ となり、過小評価となる確率 $P_I = 0$ である(全てか過大評価)。同様に P_{II} , P_{III} に関しても

$$P_{II} = \int_{\tilde{\Delta}_{bias} - \sqrt{\tilde{\Delta}_{bias}^2 + 1}}^{\tilde{\Delta}_{bias} + \sqrt{\tilde{\Delta}_{bias}^2 + 1}} \phi(\tilde{\gamma}) d\tilde{\gamma} \quad (29)$$

$$P_{III} = \int_0^{2\tilde{\Delta}_{bias}} \phi(\tilde{\gamma}) d\tilde{\gamma} \quad (30)$$

Fig.4 Probabilities P_I , P_{II} and P_{III} as functions of $\tilde{\Delta}_{bias}$.

と導かれる。式(27), (29)および(30)を数値的に積分して得られた P_I , P_{II} および P_{III} を $\tilde{\Delta}_{bias}$ の関数として図 4 に示す。

図 4 より, $\tilde{\Delta}_{bias} \gg 1$ となる領域 ($\Delta_{bias} \gg \sigma_{ref}$, $\Delta_{bias} \gg u_{ref}$) では, $P_I = P_{II} = 0.5$ となることからわかる。この場合, 50% の確率で不確かさの過小評価が存在することになるが, $\Delta_{bias} \gg u_{ref}$ の領域では, $u_I(d_{nc}) = u_{II}(d_{nc}) = D_{bias}$ と見なせる。つまり, この領域では, 不確かさ u_{ref} に起因する影響は殆どなく, 相対的に過小評価 (もしくは過大評価) の度合いが小さくなるので, 式(23) (もしくは式(8)) により, ほとんどの事例において妥当な不確かさが与えられると考えられる。また, 図 4 より, P_I は Δ_{bias}/u_{ref} が 2~3 程度までは急激な変化 (増加) を示しているが, それ以降は飽和傾向となり, P_{II} との差も小さくなっている。つまり, GUM の考え方に沿った評価式である式(23)によって与えられる不確かさにおける過小評価の度合いは, Δ_{bias}/u_{ref} が 2~3 以上であれば, 従来の評価法と大きな差はないと考えられる。したがって, 比較測長器の偏差 (系統誤差) が, 校正に使用する基準段差の不確かさ u_{ref} の 2~3 倍程度以上であると推定される場合は, 無用な不確かさの重複を避けるという観点から, 今回提示した評価式(23)により不確かさを評価することが妥当であると考えられる。

一方, 比較測長器の系統誤差 (かたより) が, 校正に使用する基準段差の不確かさと同程度かそれ以下であると推定される場合は, 従来の評価法では過小評価となる確率が殆どなく, $|\Delta_{bias}| < u_{ref}$ の範囲では, 過小評価となる確率 P_I は 0 である。これは言い換えると, ほとんどの場合に不確かさの過大評価が行われていることになる。一方, 今回提示した GUM の考え方に沿った評価式では, 図 4 より, $|\Delta_{bias}| < u_{ref}$ の範囲においては, 60% 以上の頻度で過小評価が存在することになる。なお, この過小評価の中には, $u_{II}(d_{nc}) = s/\sqrt{n}$ の様な明らかな過小評価 (比較測長器のかたよりの推定値が 0) も存在する。

今回新たに提案した不確かさ評価式 $u_{III}(d_{nc})$ に関しては, $\Delta_{bias} \gg u_{ref}$ の領域では $P_{III} = 0.5$ であり, また, $|\Delta_{bias}| < u_{ref}$ の領域においても常に $P_I < P_{III}$ であることから, 従来法ほど $u_{III}(d_{nc})$ が過大評価となる確率は高くない。さらに, $u_{III}(d_{nc}) = s/\sqrt{n}$ となるのは $\Delta_{bias} = \gamma$ となる時のみであり, 明らかな過小評価が与えられる確率はかなり低い。以上より, 式(24)により与えられる不確かさにより, 現実的な不確かさ評価式として, ある程度バランスのとれた評価が可能であると考えられる。

4 かたよりを補正しない不確かさ評価の例 2 : 温度測定

第 3 章では, ブロックゲージ比較測定法における比較測長器の校正を例にとり, かたよりを補正しない場合の不確かさ評価について述べたが, 温度に関連する不確かさ評価に関しても, 前節と同様のことが言える。具体的には, 2 本のブロックゲージの温度差の不確かさ $u(\delta\theta)$ および被校正ブロックゲージの 20 度からの偏差の不確かさ $u(\theta)$ が同様の扱い例として当てはまる。

2 本のブロックゲージの温度差の不確かさ $u(\delta\theta)$ を評価する場合は, 通常, 評価のみのために 2 本のゲージに温度計を貼り付け, 温度差 $\delta\theta$ の測定を多数回行う。一般的に, この測定結果は, 図 2 に示すような分布となり, かたより D_T と標準偏差 s_T を持つ。この場合本来ならば, かたより D_T 分だけ補正する, つまり $\delta\theta \neq 0$ として扱わなければならない。しかしながら, 前述した様に, 温度差測定は常に行っているわけではないので, かたよりが存在するといってもその値に多少の曖昧性があるといった理由から, 実際のブロックゲージ測定の際には, やはり $\delta\theta = 0$ として, かたより D_T は不確かさに含めてしまうことがほとんどである。

従来のかたより D_T を含めた不確かさの評価法では, 標準偏差 s_T および温度計の温度差測定の不確かさ u_T を考慮して,

$$u_I^2(\delta\theta) = D_T^2 + s_T^2 + u_T^2 \quad (31)$$

として, 温度差の標準不確かさ $u_I(\delta\theta)$ が求められている⁵⁾。しかしながら, 前節で述べた様に, 温度差測定の不確かさ u_T を加えることは明らかに厳密性を欠いている。

温度差測定の場合の構造模型を考えてみると, 前節と同様に, 測定値 $\delta\theta$ に関して

$$\delta\theta = \Delta_T - \gamma_T \quad (32)$$

の関係式が成立する。ここで, Δ_T は実際の温度差, γ_T は温度計の持つ温度差測定に対する偏差である。なお, 温度計のばらつき (偶然効果) は十分に小さいとして含まれていない。また, 前節と同様に, $E[\gamma_T] = 0$, $V[\gamma_T] = E[\gamma_T^2] = u_T^2$ の関係式が成立する。これらの関係式から, (32)式の両辺の 2 乗の期待値を計算すると,

$$E[\delta\theta^2] = \Delta_T^2 + u_T^2 \quad (33)$$

の関係式が成立する。

前節と同様に, GUM の考え方によれば, ここでの求めたい標準不確かさ $u_{II}^2(\delta\theta)$ は, 実際の温度差 Δ_T^2 の推定値である。したがって, 式(33)より, $u_{II}^2(\delta\theta)$ は

$$u_{II}^2(\delta\theta) = D_T^2 + s_T^2 - u_T^2 \quad (34)$$

と与えられる。なお, ここでは $E[\delta\theta^2]$ の推定値として, 付録 (A5) 式の関係式から, $D_T^2 + s_T^2$ が用いられている。

式(31)と(34)を比較すると, 比較測長器の系統誤差の場合と同様に, 従来の評価法の方が $2u_T^2$ だけ値が大きいことが分かる。ただし, 平均値で評価する比較測長器の系統誤差と違い, 温度差測定の場合は, 測定毎のばらつき効果 s_T を評価に加えておく必要がある。したがって, 3.5 節の様に過小評価 (過大評価) の程度を考える場合は, $\sqrt{D_T^2 + s_T^2}$ と u_T との大きさの関係により不確かさの評価法を考慮していく必要がある。

また、被校正ブロックゲージ温度の 20 度からの偏差の不確かさ $u(\theta)$ に関しても、基本的には同様の扱いとなる。この場合も、評価のみのためにブロックゲージ温度を多数回測定し、20 度からの偏差の 2 乗の期待値（平均値（かたより）とばらつきの 2 乗和）を求める。この 2 乗の期待値に温度計そのものの不確かさを考慮することにより、式(34)と同じ形で不確かさ $u(\theta)$ を推定することができる（従来法の場合は式(31)と同じ形）。ただし、ここでの温度計そのものの不確かさとは、前述の温度差測定に対する不確かさ u_T とは違い、絶対温度測定に対する不確かさであることに注意する。

5. 考 察

3 章および 4 章において、かたよりを補正しない場合の GUM の考え方に沿った不確かさの評価法について提示したが、この評価式は不確かさ要因の減算であるため、前述した様に、かたよりの推定値が 0 となってしまう場合があり得る。実際の不確かさ評価においては、0 の様な明らかな過小評価は採用しづらく、また、実用的にも問題がある場合がある。例えば JCSS 制度のように、測定能力を不確かさの大きさを用いて評価、認定しようとする状況では、公平性の点から望ましくないだけでなく、実際には測定能力が低いにもかかわらず、十分な能力を有していると認定してしまう可能性が高まることとなる。

今回例に挙げたブロックゲージ比較測長器の系統誤差評価の場合、一般的な比較測長器のかたより（系統誤差）は、通常使用される基準段差の校正不確かさと同程度かそれ以下であると推定される。この場合、図 4 に示した様に、今回提示した評価式（式(23)）では不確かさの過小評価となる頻度、度合いが大きくなる。JCSS 制度の様に、多数の事業者が同じ校正サービスを行う状況を想定すると、明らかな過小評価とそうでない事業者が混在することになる。したがって、計測の信頼性を担保するしくみとして、式(23)は必ずしも適切な評価法とは言えない面がある。さらに、式(23)においては、基準段差の校正不確かさ u_{ref} が妥当な値であることが前提となっているが、実際の校正証明書などに記載されている校正不確かさ u_{ref} は、安全を見て大きめ（過大評価）の値に設定されていることが多い。結果、比較測長器の不確かさ $u_{II}(d_{nc})$ の更なる過小評価につながる可能性が高い。このような場合は、式(8)で示される従来の不確かさ評価法も許容されるべきと考えられる。ただし、この場合は、明らかな過小評価を防ぐ代わりに、不確かさの過大評価となる可能性が高いことを認識しておく必要がある。

また、今回新しく提案した評価式（式(24)）関しても、3.5 節で述べた様に、上述した様な過小評価を避け、不確かさの安全側評価が望ましい場合には、実用上、ある程度推奨できると考えられる。ただし、この評価式も、明らかな過小評価を防ぐ代わりに、不確かさの過大評価となる可能性があることを認識しておく必要がある。また、従来法と同様に数学的な厳密性がないという点も留意しなければならない。

最後に、実際に比較測長器の系統誤差の不確かさ評価を行う際の留意点について触れておく。実際に比較測長器の系統誤差を既知の基準ブロックゲージ対で評価する場合は、複数のブロックゲージ対、つまり、値の異なる複数の段差

を用いて評価することがほとんどである。この場合は、複数の測定結果より、測定範囲内の平均的なかたよりの 2 乗を推定することになる。例えば、3 つの段差を系統誤差の評価に用い、その測定結果が D_{bias^k} ($k=1, 2, 3$) であった場合、かたよりの 2 乗の推定値 D_{bias^2} としては、3 つのデータの 2 乗の期待値を採用することにより、

$$\begin{aligned} D_{bias^2}^2 &= E[D_{bias^2}^2] \\ &= \text{ave}^2[D_{bias^2}] + \text{stdev}^2[D_{bias^2}] \end{aligned} \quad (35)$$

と与えられる。ここで、 $\text{ave}[\bullet]$ はデータの平均値を、 $\text{stdev}[\bullet]$ は標準偏差を表している。式 (35) より、求めたい不確かさ $u_1(d_{nc})$, $u_{II}(d_{nc})$, $u_{III}(d_{nc})$ は、式中の D_{bias^2} が d_{nc} , D_{bias^2} に置き換えられた形で与えられることになる。

6. 結 言

本報告では、ブロックゲージの比較測定法を例にとり、かたよりを補正しない場合の不確かさ評価法に関する考察を行った。従来のかたよりを補正しない場合の評価法においては、かたよりや基準器の不確かさといった要因が単純に全て加算されていたが、今回、GUM の考え方に沿って計算した結果、従来法とは違う評価式（不確かさの減算）が得られた。二つの評価式の比較結果を以下に要約する。

- (1) 従来の評価法では、不確かさの過大評価となってしまう可能性が高い。
- (2) 今回新しく提示した評価法では、期待値としては妥当な不確かさを与えようと考えられるが、個々の事例においては、明らかな不確かさの過小評価となってしまう場合が有り得る。
- (3) 補正しないかたよりの推定値が、校正に使用する基準器の不確かさと同程度以下の場合は、不確かさ評価法の扱いに特に留意する必要がある。

一般に不確かさの評価を行う場合、考え得る限りの要因を挙げ、それら全てを加算するなど、とにかく過小評価を防ごうとする傾向が強い。しかしながら、こういった傾向は必ずしも妥当な不確かさを評価することには繋がらない。GUM6.3.1 節においても、「不確かさの評価と、測定量に対する安全限界の付与を混同してはならない」との記述がある。一方で、確率分布を基軸とした GUM の考え方では、今回示した例の様に、採用しづらい不確かさ（明らかな過小評価）となる場合も有り得る。実用的には、個々の事例において不確かさの扱いを考慮する必要があると思われる。

付 録

A かたよりを補正する場合の不確かさの求め方

ここでは、式(5)を厳密に導出する。式(4)に、(9)、(10)、(11)を代入することにより、

$$d_c = \lambda - \varepsilon_{ref} + \gamma + \varepsilon \quad (A1)$$

ε , ε_{ref} , γ が互いに独立な確率変数であることから、

$$V[d_c] = \frac{\sigma^2}{n_{ref}} + \sigma_{ref}^2 + \frac{\sigma^2}{n} \quad (A2)$$

$u^2(d_c)$ は $V[d_c]$ の推定値ゆえ、右辺の σ^2 , σ_{ref}^2 をそれぞれの推定値 s^2 , u_{ref}^2 で置き換えることにより、式(5)を得る。

B かたよりを補正しない場合（従来法）の不確かさ（式（7））の求め方

一般に，かたより μ_x ，分散 σ_x^2 を有する確率変数 x があるとする．すなわち，

$$E[x] = \mu_x \quad (A3)$$

$$V[x] = \sigma_x^2 \quad (A4)$$

このとき，

$$\begin{aligned} E[x^2] &= E[\{\mu_x + (x - \mu_x)\}^2] \\ &= \mu_x^2 + 2\mu_x E[x - \mu_x] + E[(x - \mu_x)^2] \\ &= \mu_x^2 + \sigma_x^2 \end{aligned} \quad (A5)$$

式(7)において， $u_1^2(d_{nc})$ は $E[\varepsilon_{nc}^2]$ の推定値と考えられるから，上式で μ_x^2 の推定値として D_{bias}^2 を， σ_x^2 の推定値として $u^2(d_c)$ を用いることにより，式(7)一行目の式を得る．

参 考 文 献

- 1) BIPM, IEC, IFCC, ISO, IUPAC, OIML, Guide to the expression of uncertainty in measurement, ISO (1993).
- 2) 飯塚幸三監修：計測における不確かさの表現のガイド，日本規格協会，(1996).
- 3) J. E. Decker and J. R. Pekelsky: Uncertainty of gauge block calibration by mechanical comparison, NRC Doc. No. 39998 (1996).
- 4) 首藤俊二：ブロックゲージ比較測定における不確かさ，精密工学会誌，**65**，7 (1999) 953.
- 5) 製品評価技術基盤機構認定センター：JCSS 不確かさ見積もりに関するガイド(ブロックゲージ)，第2版，(2005).