Revision of ISO 13528:
Statistical Methods for Proficiency Testing by Interlaboratory Comparison

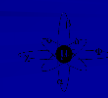
8th Eurachem PT Workshop

7 October, 2014

Daniel Tholen, M.S.
Traverse City, MI

Revision of ISO 13528

- ISO TC69
  - SC6 WG9
- Revision process
  - CD & DIS Ballots
  - FDIS
- Major changes from current version



INTERNATIONAL STANDARD **ISO FDIS 13528**

Final Draft International Standard
2014-09-18

Statistical methods for use in proficiency testing by interlaboratory comparison

Mèthods statistiques utilisèes dans les essais d'aptitude

ISO TC69: *Application of Statistical Methods*

- TC69 has a broad scope:
  - SC1: Terminology and symbols
  - SC4: Process management (control charts)
  - SC5: Acceptance sampling
  - SC6: Measurement methods and results
  - SC7: Methods to support Six Sigma
  - SC8: New technology & product development
  - AHG1: Documents to support ISO/IEC 17025 and ISO 15189
  - WG3: Interpretation of data
- Secretariate: ANSI (USA)...held at ASQ

ISO TC69: *Application of Statistical Methods*

- TC69 is a "Horizontal Committee" in ISO – all other TCs are obliged to follow TC69 Standards, if applicable
  - TC176 (ISO 9001) is only other Horizontal C.
- TC69 Standards must therefore be broadly applicable – no language for specific use
  - 97 Standards
  - 19 Participating member countries
  - 31 Observing members
  - 12 Organizations in liaison (including ILAC)

By Comparison with ISO TC69

- ISO CASCO (ISO/IEC 17043, 17025, etc)
  - 80 Participating members
  - 43 Observing members
  - 19 Organizations in Liaison (including ILAC)
  - 33 Standards
- ISO TC 176 (Quality Management)
  - 95 Participating members
  - 25 Observing members
  - 24 Organizations in Liaison (including ILAC)
  - 3 Standards

ISO TC69/SC6

- TC69 Subcommittee 6 on *Measurement methods and results*
  - Secretariat is JISC (Japan)
  - 31 Standards
  - ISO 5725: Accuracy (trueness and precision)
  - ISO 11843: Capability of detection
  - Standards related to calibration
  - Standards related to measurement uncertainty
- TC69/SC6 Working Group 9 on *Statistical methods for proficiency testing*

ISO 13528:2005

- Project proposed in 1997; published 2005
- Written as a requirements document
  - Statistically optimal methods for evaluating bias and repeatability
- Widely adopted (in part) by ILAC member countries
- ISO 13528:2005 reaffirmed in 2009

Revision of ISO 13528

- ILAC work proposal to revise ISO 13528 was approved in 2010
  - Update to accommodate ISO/IEC 17043
  - Fix errors, remove obsolete procedures
  - Add a few new procedures
- "Limited revision" approved by TC69/SC6
- CD1 ballot closed March, 2011
  - Approved: 14 'yes' 4 'No'
  - 600+ comments, rewrite needed
  - (status at 7th Eurachem PT Workshop)

DIS1 and DIS2 ISO 13528

- CD2 ballot (2012) not conducted
  - WG reviewed draft rewrite as DIS
  - Rearrange text and examples
- DIS1 ballot closed April, 2013
  - Approved: 14 'Yes' 3 'No'
  - 800+ comments, further rewrite needed
  - ISO granted 1 year extension for development
- DIS2 ballot closed, February, 2014
  - Approved: 14 'Yes' 1 'No'
  - 800+ comments, many changes

FDIS 13528

- TC69/SC6/WG9 met in June, 2014
  - Agreed to many changes
  - Agreed to not make significant changes that required re-ballot as DIS
  - Agreed to ballot as FDIS
- Draft FDIS to WG in July, many useful edits
- Will send to ISO editors by Oct 15
- Ballot by January, 2015 (2 months)
- If successful, publish in June, 2015

What will change - general

- Accommodate new requirements in ISO/IEC 17043
  - Design
  - Qualitative analytes
  - Scope (inspection, sampling, individual)
- Rearranged text and examples
- More guidance, fewer requirements
  - Different ILAC regions and countries will use ISO 13528 differently

What will not change

- Algorithm A (except with minor change to stopping criterion)
- Homogeneity testing procedure
- Stability testing procedure
- Techniques to determine assigned value and SDPA
- Allowance for other statistically valid procedures

10. Invalid/obsolete techniques removed

- Requirements on replicates and rounding
  - Often different than routine testing
- CUSUM across rounds
- Evaluation by rank
- Youden plot ellipses
  - Youden plot retained
- Requirement to not allow ‘<‘ values

9. New Symbols

- Assigned Value: Current X ➔ new x_{pt}
- Standard Deviation for Proficiency Assessment (SDPA): Current $\hat{\sigma}$ ➔ new σ_{pt}
- Different than ISO/IEC 17043 ☹

9. Example of New Symbols: z Score

Current: $z = \frac{(x_i - X)}{\hat{\sigma}}$

New: $z = \frac{(x_i - x_{pt})}{\sigma_{pt}}$

8:Difficult issues raised, but not fully resolved

- Handling censored values (<)
- Calculating x_{pt} and u_{char} when results are from 2+ experts and include u_{lab}
  - Procedure in current document is incorrect
  - PT provider to determine what they do
- Using PT data to estimate method repeatability σ_r and reproducibility σ_R

7. Guidance on evaluation with Total Measurement Error

- D & $D\%$ as performance statistics

  $D_i = (x_i - x_{pt})$ $D_i\% = D_i / x_{pt}$ x 100 %
- δ_E as criterion for measurement error, as units or as % of x_{pt}

  Usually $\delta_E = 3\sigma_{pt}$
- P_A as standardized score (proportion of allowed error)

  $P_A = D_i / \delta_E$ or $P_A = D_i\% / \delta_E$

6. Can use z' with consensus mean and standard deviation

- Correlation between x_i and x^* and s^* is usually small, so no correction is needed.

$$z' = \frac{(z_i - x_{pt})}{\sqrt{\sigma_{pt}^2 + u^2(x_{pt})}}$$

☺

5. Align with ISO Guide 35 (rev)

- Usually, proficiency test items are a type of reference material
- Similar concerns for homogeneity, stability, assigned values
  - Allow use of experience on previous batches to inform on homogeneity and stability
  - With experience, can have limited tests
- PT and RM should have a common approach to uncertainty of assigned values

5: Guide 35 components for uncertainty of the assigned value

$$x_{pt} = x_{char} + \delta_{hom} + \delta_{trans} + \delta_{stab}$$

with $x_{char} = x^*$ or x_{ref} or x_{CRM}

$$u(x_{pt}) = \sqrt{u_{char}^2 + u_{hom}^2 + u_{trans}^2 + u_{stab}^2}$$

5: Additional components for uncertainty of the assigned value

- Primarily needed when reference values are used as x_{pt} and σ_{pt}
- When consensus values are used, u_{char} may be sufficient as $u(x_{pt})$
  - Much of the uncertainty due to inhomogeneity and instability will be included in σ_{pt}
- Often, components for transport stability and long term stability will be zero

4: Add new Robust techniques

- Simple:
  - median for x_{pt}
  - $MADe$ and $nIQR$ for σ_{pt}
- Computationally intense:
  - Hampel estimate for x_{pt}
  - Q and Qn for σ_{pt}
  - From Rousseeuw, Uhlig, Hampel, ISO/TS 20612

3: Flag questionable participant uncertainties (u_{lab})

- Similar to IMEP practice

$u_{min} = u_{ref}$ flag $u_{lab} < u_{min}$

$u_{max} = 1{,}5s^*$ flag $u_{lab} > u_{max}$

When no u_{ref}, PT provider chooses u_{min}

2: New Statistical Model

Current: $x_i = \mu + B_i + \varepsilon_i$ (from ISO 5725)

New: $x_i = \mu + \varepsilon_i$

With: x_i = Participant result, lab i

μ = True value for measurand

B_i = Laboratory Bias, lab i

ε_i = Random error, lab i

2: New Statistical Model

- Current model leads to design requirements that are largely ignored
  - replicates, rounding
- Change statistical model to remove estimation of laboratory bias
- Performance evaluation is evaluation of the fitness of the submitted result, not evaluation of an estimate of Bias

1: Link Statistical Methods to objectives for the PT scheme

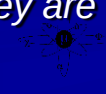
- FDIS ISO 13528:

4.1.2 *The statistical design and data analysis techniques shall be consistent with the stated objectives for the proficiency testing scheme.*

- Different objectives could require different statistical methods

ISO/IEC 17043 clause 4.4

- **4.4.1.3** *The proficiency testing provider shall document a plan … that addresses the objectives, purpose and basic design…*
- **4.4.4.1** *Statistical designs shall be developed to meet the objectives of the scheme…*
- **4.4.4.2** *The proficiency testing provider shall document the statistical design and data analysis methods to be used … and shall provide a description of the reasons for their selection and assumptions upon which they are based.*



1: Link Statistical Methods to objectives for the PT scheme

- Frequently stated objective: "Objective is to determine the competence of participants to measure….(*some measurand in something*)"
- How do you define 'competence'?
  - "Agreement with the true result, close enough to allow a correct decision"
  - "Agreement with other participants, with typical measurement error"
  - "Confirm claims for uncertainty"

1: Link Statistical Methods to objectives for the PT scheme

- What do we compare a participant's results to?
  - Reference values and fitness criteria; or
  - Results from other participants;
  - The participant's claim for uncertainty
- Could be mixed for x_{pt} and σ_{pt}
- Could have more than one score
  - Recommend: use z and ζ

1: Link Statistical Methods to objectives for the PT scheme

- Other purposes:
  - Assess comparability of measurement methods
  - Investigate reasons for errors
  - Satisfy regulatory needs
- More choices in statistical methods ➔ Need better description of objectives

Message from new ISO 13528

- Use statistical techniques appropriate for objectives, type of data, number of results
- Use technical expertise and judgment
  - experience with interlaboratory comparisons
  - understanding of statistical methods
- Cannot rely blindly on formulae