

MÉTODOS DE EVALUACIÓN POR EL PROVEEDOR DE LA CALIDAD DE LOS RESULTADOS DE UN EJERCICIO DE INTERCOMPARACIÓN ACÚSTICA

PACS: 43.15.+s

Espinel Valdivieso, Ana¹; Arenaz Gombau, Ángel¹; Ausejo Prieto, Miguel¹; Gómez García, Sonsoles²

¹ AUDIOTEC Ingeniería Acústica, S.A.

C/ Juanelo Turriano, nº 4 - Parque Tecnológico de Boecillo - 47151 Boecillo (Valladolid), España

Tel: +34 983 361 326, E-mail: info@audiotec.es

² Asociación Española para la Calidad Acústica (AECOR)

C/ Tambre, nº 21, 2º - 28002 Madrid (Madrid), España

Tel: +34 915 541 421, Fax: +34 915 546 796, E-mail: info@aecor.es

ABSTRACT

UNE EN ISO/IEC 17043:2010 specifies general requirements for the competence of providers of proficiency testing schemes and for the development and operation of proficiency testing schemes, dedicating a specific section to data analysis and evaluation of proficiency testing scheme results. In this work we will present the method used by the provider for the evaluation of acoustics proficiency testing scheme results by meeting the quality requirements of the standard.

RESUMEN

La norma UNE EN ISO 17043 especifica los requisitos generales para la competencia de los proveedores de ensayos de intercomparación y para el desarrollo y operación de ensayos de intercomparación, dedicando un apartado específico al análisis de datos y a la evaluación de los resultados de los ensayos de aptitud. En este trabajo, se va a tratar la sistemática empleada por el proveedor para la evaluación de la calidad de los resultados en un ejercicio de intercomparación en acústica cumpliendo los requisitos de calidad establecidos en la norma de aplicación.

1. INTRODUCCIÓN

La norma UNE EN ISO/IEC 17043 especifica los requisitos generales para la competencia de los proveedores de ensayos de intercomparación y para el desarrollo y operación de ensayos de intercomparación, dividiéndolos en requisitos técnicos y requisitos de gestión, dedicando un apartado específico al análisis de datos y a la evaluación de los resultados de los ensayos de aptitud [1].

La Asociación Española para la Calidad Acústica (AECOR) y la empresa Audiotec trabajan conjuntamente desde hace diez años en la promoción, organización y el correcto desarrollo de ejercicios de intercomparación de laboratorios de acústica, tanto en materia ambiental como en edificación, habiendo promovido desde entonces nueve ejercicios de intercomparación en los que se han inscrito más de 200 participantes, habiéndose consolidado como el proveedor líder y de referencia a nivel nacional en el campo de la acústica.

En este trabajo, se va a tratar la sistemática empleada por el proveedor para la evaluación de la calidad de los resultados en un ejercicio de intercomparación en acústica. Como se verá más adelante, es imprescindible que previamente al análisis de los resultados, el proveedor garantice que el desarrollo de la intercomparación se haya llevado de forma adecuada y cumpliendo los requisitos de calidad establecidos en la norma antes mencionada.

2. CONTROL DE CALIDAD DEL DESARROLLO DEL EJERCICIO DE INTERCOMPARACIÓN PREVIO A LA EVALUACIÓN DE LOS RESULTADOS

Durante el desarrollo del ejercicio de intercomparación acústica, todos los laboratorios participantes acuden a las mismas instalaciones a realizar sus ensayos sobre el mismo ítem. Con el fin de asegurar que los ítems bajo ensayo permanecen inalterables durante el desarrollo del programa, el proveedor debe realizar periódicamente controles de estabilidad y homogeneidad, asegurando así que se mantienen las condiciones iniciales de los ítems e instalaciones para todos los laboratorios.

El proveedor, además, debe elaborar un protocolo específico para cada tipo de ensayo que indique una serie de pautas a seguir por los participantes con el fin ayudarles en la correcta ejecución del ejercicio y de minimizar los errores más comunes que puedan cometer durante su participación en la intercomparación. Estos errores comunes que cometen los participantes son los siguientes:

- No seguir las instrucciones del organizador
- No realizar el ensayo de forma habitual sino como un ítem especial
- Modificar las indicaciones del protocolo con el fin de realizar el ensayo en condiciones más favorables
- No respetar las condiciones de repetibilidad, mezclando metodologías, personal, equipos, etc.
- Realizar un mayor número de medidas del requerido por el organizador o el método de medida, seleccionándolas o promediándolas en virtud de los resultados obtenidos.
- Cometer errores en la transcripción de resultados.
- Comentar resultados con otros participantes del ejercicio
- Presentar como propios resultados de medidas subcontratadas a otros laboratorios

Para que el desarrollo de un ejercicio de intercomparación se considere que ha tenido la calidad mínima requerida, el proveedor debe asegurar y comprobar que:

- La organización ha sido adecuada y se han suministrado las instrucciones precisas y adecuadas para ayudar a los laboratorios en la correcta realización del ejercicio.

- Se ha dispuesto de ítems de ensayo e infraestructuras adecuadas para la realización de cada tipo de ensayo objeto de intercomparación, y que éstas han permanecido inalterables y estables en el tiempo durante todo el desarrollo del programa, habiéndose realizado controles periódicos de estabilidad y homogeneidad.
- No se han observado por parte del técnico de logística circunstancias que pudieran haber puesto en compromiso el que los laboratorios contaran con las mismas condiciones de ensayo.
- El número de laboratorios participantes y repeticiones ha sido adecuado.
- Los laboratorios han realizado los ensayos con equipos de ensayo calibrados y cumpliendo los requisitos normativos exigidos.
- Los laboratorios participantes han enviado los datos correctamente en los formatos elaborados por el proveedor y en los plazos acordados.
- Se pueda considerar que los laboratorios han empleado bien los protocolos enviados.
- No se han recibido quejas o reclamaciones por parte de los participantes desde el inicio del programa hasta la fecha de emisión del informe.

Una vez que el proveedor ha asegurado y comprobado que se han cumplido todos los requisitos mencionados y disponga de los resultados de todos los laboratorios que han participado en el ejercicio, se procede a procesar los resultados y a evaluar su calidad en base a los requisitos establecidos en la norma de aplicación.

3. ANÁLISIS DE DATOS Y EVALUACIÓN DE LA CALIDAD DE LOS RESULTADOS DEL EJERCICIO DE INTERCOMPARACIÓN

Una vez que todos los laboratorios han enviado los resultados de sus ensayos al proveedor, éste lleva a cabo su análisis y evaluación para la emisión de un informe final.

3.1. Análisis de datos y registros

Antes de comenzar con el análisis de datos y registros, el proveedor debe validar todo el equipo de procesamiento de datos y asegurar el correcto almacenamiento de la información. El mantenimiento del sistema informático deberá incluir un proceso de copia de seguridad y un plan de recuperación del sistema. Los resultados de tal mantenimiento y de las verificaciones operativas deberán quedar registrados.

Además, el proveedor debe establecer una sistemática para validar y chequear los datos recibidos por los participantes antes de proceder a realizar su análisis estadístico. En primer lugar, a la recepción de los datos enviados por los laboratorios, se comprobará que no haya errores de bulto tales como errores de posicionamiento de comas en los resultados con decimales, que en alguna de las cifras se haya introducido un número a mayores o que se intuya que ha habido un intercambio de cifras entre celdas. En estos casos, el proveedor contactará con el laboratorio con el fin de que le envíe la hoja de cálculo corregida. En caso de que no la envíen, el proveedor podrá tomar la decisión de eliminar dichos valores.

Una vez descartados los errores de bulto, el proveedor debe introducir los resultados de todos los participantes en la hoja de cálculo diseñada para el tratamiento estadístico con el código correspondiente a cada laboratorio. Para verificar la validez de la transferencia de los datos, primero se introducirán los resultados de todos los laboratorios en una pestaña y, una vez se hayan introducido todos los resultados, se repetirá el mismo procedimiento en una segunda pestaña. Tras introducir los resultados en las dos pestañas, se revisará una tercera pestaña de chequeo de datos para comprobar que coinciden los resultados de cada laboratorio en ambas pestañas. En el caso de que todos los resultados introducidos en las dos primeras pestañas coincidan uno a uno, aparecerá un mensaje indicando que la transferencia de datos ha sido

correcta y, por lo tanto, el proveedor podrá seguir adelante con el tratamiento estadístico de los datos.

El análisis estadístico llevado a cabo en la hoja de cálculo a partir de los resultados aportados por los participantes, deberá generar un resumen estadístico, estadísticas de desempeño e información coherente con el diseño estadístico del programa de ensayos de aptitud. Será necesario que el proveedor establezca un procedimiento de validación que asegure la validez de los cálculos realizados en cada una de las hojas de cálculo diseñadas. En este caso, las validaciones de cada una de las pestañas de cálculo se realizarán de forma manual y de forma automatizada en nuevas hojas de cálculo destinadas a tal fin.

Con el fin de minimizar la influencia de los valores atípicos en el resumen estadístico, el proveedor utilizará en el tratamiento de datos métodos estadísticos robustos o pruebas apropiadas para detectar valores estadísticos atípicos como los test de Cochran o de Grubbs.

3.2. Obtención de datos estadísticos para poder realizar la evaluación del desempeño

Con el fin de satisfacer el propósito del programa, el proveedor de ensayos de aptitud debe utilizar métodos de evaluación válidos. Estos métodos se deben documentar y deben incluir una descripción de las bases para la evaluación por parte de los laboratorios. Además, los datos tendrán que ser tratados por un experto estadístico con experiencia acreditada en el tratamiento de datos acústicos, lo cual facilitará la comprensión de los resultados obtenidos.

La evaluación del desempeño de los participantes en un ejercicio de intercomparación acústica, se lleva a cabo empleando los procedimientos establecidos en la norma ISO 13528:2005 [2]. En esta norma se detallan todos los parámetros estadísticos y cálculos incluyendo:

- Cálculos estadísticos para el diseño y la interpretación de los ensayos de aptitud.
- La presentación de descripciones detalladas de métodos estadísticos que se utilizan en el análisis de los datos obtenidos de los programas de ensayos de aptitud.
- La presentación de recomendaciones sobre su utilización en la práctica.

A partir de las indicaciones de esta norma y tomando como referencia el documento G-ENAC-14 [3], una vez calculados los valores promedio, la desviación y típica y la varianza de los resultados aportados por cada laboratorio, se determinan los siguientes parámetros:

- Valor asignado (x): obtenido según el método de consenso entre los laboratorios participantes, correspondiéndose al valor promedio robusto de los resultados obtenidos por los laboratorios participantes de acuerdo al análisis robusto del algoritmo A del anejo C de la norma ISO 13528:2005.
- Desviación estándar diana u objetivo (σ_p): obtenida mediante el promedio de las desviaciones estándar calculadas en ejercicios de intercomparación anteriores organizados por el proveedor o bien a partir de lo indicado en la norma ISO 12999-1:2014.
- Desviación estándar del ejercicio ($\sigma_{ejercicio}$): obtenida mediante la desviación estándar que describe la dispersión estándar de resultados de los participantes en el ejercicio de intercomparación.
- Incertidumbre del valor asignado (u_x): calculada a partir de la desviación estándar del ejercicio y del número de laboratorios participantes.

- Z-score: parámetro para la evaluación del rendimiento de los laboratorios, calculado a partir de la media de cada laboratorio participante, el valor asignado por consenso y la desviación estándar diana u objetivo.

3.3. Evaluación de la calidad de los resultados del ejercicio de intercomparación

Con el fin de ayudar a la interpretación y evaluación de los resultados del ensayo de aptitud, es necesario transformarlos en un estadístico de desempeño. De este modo, será posible medir la desviación del resultado de ensayo de cada participante con el valor asignado y evaluarla con criterios de desempeño.

Es este caso, para la evaluación del desempeño se utilizará el parámetro estadístico “z-score” comentado en el apartado anterior, que se define de acuerdo con la ecuación:

$$z = \frac{x - x_a}{\sigma_p}$$

Donde:

- x_a es el valor asignado por consenso (la media robusta de los resultados remitidos por los laboratorios)
- σ_p es la desviación estándar diana o adecuada al fin pretendido
- x es la medida del laboratorio participante

El criterio de interpretación empleado para la evaluación de los resultados obtenidos por los laboratorios en este estadístico es el siguiente:

- $|Z| \leq 2$ = satisfactorio
- $2 < |Z| \leq 3$ = cuestionable
- $|Z| > 3$ = no satisfactorio

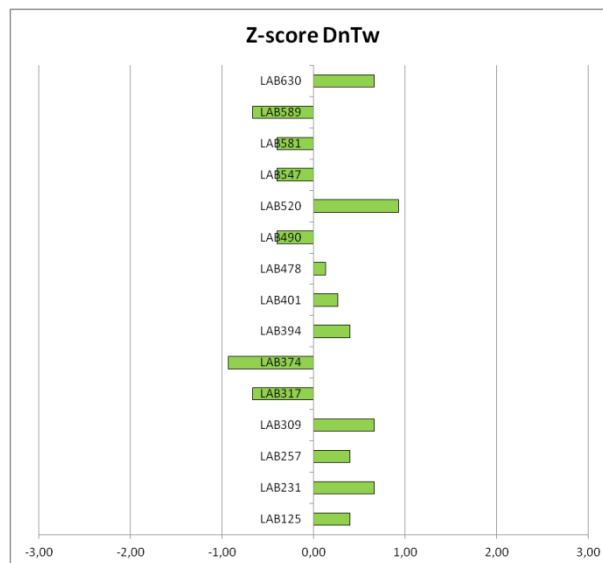


Figura 1 –Ejemplo de z-score satisfactorio

Cuando esta relación supere el valor de 0,5, el proveedor deberá informar a los participantes de que la evaluación realizada debe considerarse solamente a título informativo.

- Histogramas:

La representación gráfica en forma de histograma de los datos o parámetros de evaluación del rendimiento ayudará al proveedor a visualizar la distribución de los resultados obtenidos y le facilitará la interpretación de algunos aspectos difícilmente evaluables sólo mediante datos numéricos como, por ejemplo, la normalidad de la distribución.

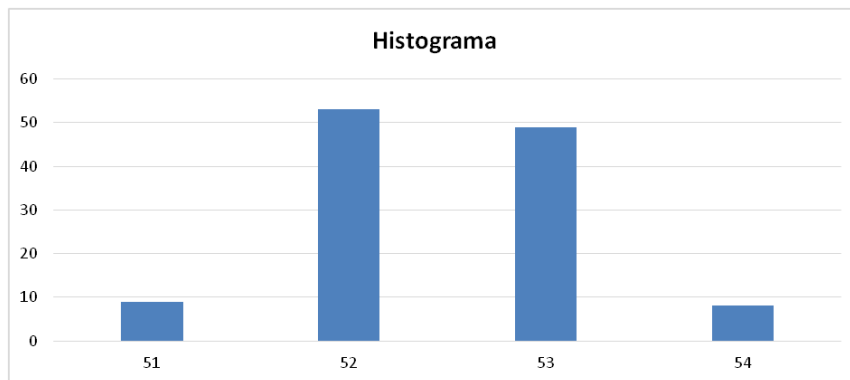


Figura 3 – Ejemplo de histograma con distribución unimodal

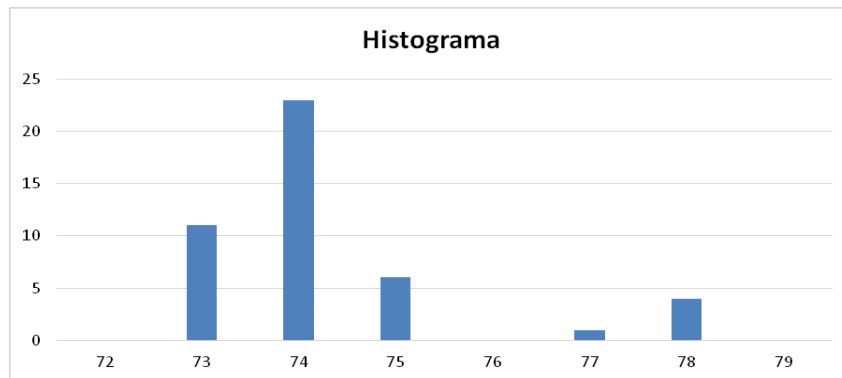


Figura 4 – Ejemplo de histograma con distribución bimodal

3.4. Casos de evaluación de la calidad de los resultados de un ejercicio de intercomparación

3.4.1. Caso 1: Evaluación satisfactoria

El proveedor considerará que la calidad de los resultados obtenidos de un ejercicio de intercomparación es satisfactoria cuando todos los valores del “z-score” sean satisfactorios, es decir, estén comprendidos entre -2 y 2, y el resto de indicadores cumplan los requisitos de calidad indicados, es decir, que la relación entre la desviación estándar del ejercicio y la desviación estándar diana u objetivo sea menor que 1.2 y que la relación entre la incertidumbre del valor asignado y la desviación estándar diana u objetivo sea menor o igual que 1.

3.4.2. Caso 2: Evaluación no satisfactoria

Se puede dar el caso de que los valores del “z-score” sean satisfactorios, es decir, se encuentren entre -2 y 2, pero el resto de indicadores no cumplan con los requisitos de calidad establecidos, es decir, que la relación entre la desviación estándar del ejercicio y la desviación estándar diana u objetivo sea mayor que 1.2 y/o que la relación entre la incertidumbre del valor asignado y la desviación estándar diana u objetivo sea mayor que 1. Esto puede ocurrir porque exista una gran dispersión de los resultados o porque los resultados se distribuyan de forma no unimodal. En estos casos, el proveedor considerará que la calidad de los resultados obtenidos de un ejercicio de intercomparación no es satisfactoria.

4. BIBLIOGRAFÍA

- [1] **UNE-EN ISO/IEC 17043:2010**. Evaluación de la conformidad. Requisitos generales para los ensayos de aptitud (ISO/IEC 17043:2010).
- [2] **ISO 13528:2005**. Statistical methods for use in proficiency testing by interlaboratory comparisons.
- [3] **G-ENAC-14 Rev. 1 Septiembre 2008**. Guía sobre la participación en programas de intercomparaciones.
- [4] **NT-03 Rev.5 Septiembre 2012**. Política ENAC sobre Intercomparaciones