

**ONARC**  
**ÓRGANO NACIONAL DE ACREDITACIÓN**  
REPÚBLICA DE CUBA



**ENSAYOS DE APTITUD**  
**Metodología de trabajo para la**  
**organización y desarrollo de**  
**comparaciones interlaboratorios.**

2007  
Rev.01

## Introducción

Las comparaciones interlaboratorios se llevan a cabo por numerosos objetivos pudiendo ser usados por los laboratorios participantes y otras partes para:

- a) determinar el desempeño de laboratorios individuales para efectuar ensayos o mediciones específicos y hacer el seguimiento del desempeño continuo de dichos laboratorios;
- b) identificar problemas en los laboratorios e iniciar acciones correctivas que puedan estar relacionadas, por ejemplo, con el desempeño del personal o con la calibración del instrumental;
- c) establecer la efectividad y el grado de comparación de nuevos métodos de ensayo o de medición y, en forma similar hacer el seguimiento de los métodos establecidos;
- d) proveer confianza adicional a los clientes de los laboratorios;
- e) identificar diferencias interlaboratorios;
- f) determinar las características de desempeño de un método, a menudo conocido como 'ensayos colaborativos' tales como repetibilidad, reproducibilidad e incertidumbre;
- g) asignar valores a materiales de referencia (MR) y evaluar su adecuación al uso en ensayo específico o procedimiento de medición.

Esta metodología está dirigida fundamentalmente a la operación de comparaciones interlaboratorios para ensayos de aptitud, es decir al fin indicado en a) no obstante la mayoría de los principios y directrices indicados son aplicables a la realización de comparaciones interlaboratorio para otros fines indicados anteriormente.

Participar en programas de ensayos de aptitud provee a los laboratorios de un medio objetivo para evaluar y demostrar la confiabilidad de los datos que ellos producen. La mayoría de los programas de ensayo de aptitud comparte como característica común, la comparación de los resultados de ensayos o medición obtenidos por dos o más laboratorios. Uno de los principales usos de programas de ensayos de aptitud es evaluar la habilidad de los laboratorios para efectuar ensayos en forma competente. Esto puede incluir evaluaciones por los mismos laboratorios, por sus clientes o por otras partes, tales como los organismos de acreditación o los regulatorios. Por lo tanto, suplementa los procedimientos internos de control de la calidad de los laboratorios, al proporcionar una medida externa de su capacidad de ensayo. Tener la confianza que un laboratorio de ensayo o la calibración obtiene resultados confiables en forma consistente es de importancia primordial para los usuarios de los servicios del laboratorio.

El examen de los resultados de la participación de un laboratorio en ensayos de aptitud, sólo puede considerarse como información proporcionada sobre la competencia técnica del laboratorio de ensayo en un momento dado bajo las condiciones específicas del ensayo (o ensayos) involucrados en un programa particular de ensayos de aptitud. Los programas de ensayos de aptitud pueden ser llevados a cabo, ya sea por una organismos de acreditación de laboratorios o por otras organizaciones. Es de vital importancia que los programas de ensayos de aptitud usados por los organismos de acreditación sean operados en forma competente, efectiva e imparcial. Se deberá reconocer, sin embargo, que los organismos de acreditación de laboratorios y sus evaluadores pueden tomar en consideración la adecuación de los resultados de ensayos producidos mediante otras actividades aparte de los programas de ensayos de aptitud. Esto incluye las muestras de control,

comparación de los datos obtenidos por otros laboratorios sobre muestras divididas, desempeño de los ensayos de auditoría con materiales de referencia certificados, etc.

El ensayo de aptitud es un proceso para chequear el desempeño actual de un laboratorio y es un requerimiento para la acreditación, según lo declarado en el documento Criterios de Acreditación ONARC

El objetivo de este documento va dirigido a establecer una metodología general para la correcta organización y desarrollo de ensayos de aptitud (en lo adelante EA). La misma consta de dos partes, una de las cuales establece los requisitos técnicos que un organismo coordinador y cualquiera de sus colaboradores asociados deben cumplir para demostrar que son competentes técnicamente para proveer tipos específicos de programas de EA; la otra parte considera diferentes métodos estadísticos que permiten realizar un tratamiento adecuado de los datos provenientes de los programas ejecutados.

En general esta metodología sustenta el proceder para aplicar las NC-ISO/IEC Guías 43-1, 43-2:00 y la NC-ISO/IEC 17025:06 en nuestro país.

Para la organización y desarrollo de las comparaciones interlaboratorios se deben aplicar los principios y procedimientos establecidos en:

- a) NC-ISO/IEC Guía 43-1:2000 ‘Ensayos de aptitud por comparaciones interlaboratorios. Parte 1. Desarrollo y funcionamiento de programas de ensayos de aptitud’ que define los principios esenciales que tienen que ser considerados cuando se organizan tales comparaciones y describe los factores que deberían tomarse en cuenta para la organización y conducción de programas de ensayos de aptitud
- b) El documento **“Requisitos técnicos para la competencia de organismos coordinadores de programas de ensayos de aptitud”**, que considera los lineamientos internacionales establecidos por la ILAC al respecto y asumimos que los requisitos de gestión no era necesario referirlos, teniendo en cuenta que el ONARC dispone de un sistema de gestión de la calidad que permite el cumplimiento de los mismos según lo establecido internacionalmente al respecto
- c) El documento **“Métodos estadísticos utilizados en los ensayos de aptitud mediante comparaciones interlaboratorios”**; elaborado tomando en consideración la amplia gama de documentos normativos internacionales que contienen lineamientos estadísticos afines. El mismo complementa a lo establecido en la Guía 43-1 de la ISO al contemplar:
  - lineamientos 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;

Estos métodos estadísticos nos permiten determinar el desempeño de laboratorios individuales en la realización de ensayos o mediciones específicas y monitorear el desempeño continuo de los laboratorios, entre estos se encuentran:

- Homogeneidad y estabilidad de las muestras;
- Determinación del valor asignado y su incertidumbre típica
- Cálculo de los estadísticos de desempeño
- Determinación de la desviación típica
  
- para la evaluación de aptitud
- Métodos gráficos para combinar los valores de desempeño tras varias rondas de un programa de ensayos de aptitud.

También se da descripción detallada y el procedimiento a seguir en los programas de ensayos de aptitud, siguientes:

- de muestras divididas
- de valores conocidos
- de comparaciones de mediciones
- de ensayos interlaboratorio

Los mencionados documentos se encuentran en los siguientes adjuntos:

Anexo A:	“Requisitos técnicos para la competencia de organismos coordinadores de programas de ensayos de aptitud”.	Página 1
Anexo B:	“Métodos estadísticos utilizados en los ensayos de aptitud mediante comparaciones interlaboratorios”	Página 11

## Anexo A:

### REQUISITOS TÉCNICOS PARA LA COMPETENCIA DE LOS ORGANISMOS COORDINADORES DE PROGRAMAS DE ENSAYOS DE APTITUD.

INDICE Contenido	Pág.
1. Objetivo .....	2
2. Gestión, personal y entrenamiento .....	2
3. Organización y diseño .....	2
3.1 Planeamiento .....	2
<b>3.2 Preparación de las unidades de ensayo .....</b>	<b>4</b>
3.3 Homogeneidad y ensayo de estabilidad .....	4
3.4 Diseño estadístico .....	4
4. Selección del método o procedimiento .....	5
5. Conducción de un ensayo de aptitud .....	5
5.1 Instrucciones a los participantes .....	5
5.2 Manipulación y almacenamiento de los materiales .....	6
5.3 Embalaje, marcado y distribución .....	6
6. Análisis de datos e interpretación de los resultados .....	6
6.1 Análisis de datos y registros .....	6
6.2 Evaluación del desempeño .....	7
6.3 Informes de los programas de ensayos de aptitud .....	7
7 Comunicación con los participantes .....	8
8 Confidencialidad .....	9
9 Colusión y falsificación de resultados .....	9
Bibliografía .....	9

## REQUISITOS TÉCNICOS PARA LA COMPETENCIA DE LOS ORGANISMOS COORDINADORES DE PROGRAMAS DE ENSAYOS DE APTITUD.

### **1. Objetivo**

Este documento establece los requisitos técnicos que un coordinador y cualquiera de sus colaboradores asociados deben cumplir para demostrar que son competentes técnicamente para proveer tipos específicos de programas de ensayos de aptitud.

### **2. Gestión, personal y entrenamiento**

- 2.1** La coordinación y conducción del ensayo de aptitud serán llevadas a cabo sólo por coordinadores y colaboradores con experiencia en el tipo particular de unidad o material de ensayo. Los coordinadores o colaboradores asociados tendrán también competencia en la medición o ensayo de las características a ser determinadas.
- 2.2** El coordinador y los colaboradores asociados tendrán personal de dirección con la autoridad, recursos y competencia técnica necesaria para desempeñar sus tareas.
- 2.3** La gerencia del coordinador definirá los niveles mínimos de calificación y experiencia necesaria para los puestos claves dentro de la organización.
- 2.4** El coordinador tendrá suficiente personal con la necesaria educación, entrenamiento, conocimientos técnicos y experiencias para las funciones asignadas.
- 2.5** El coordinador asegurará que el personal reciba entrenamiento adicional, cuando sea necesario, para asegurar el desempeño competente de las mediciones, operación de equipos y otras actividades que afecten la calidad. Cuando sea posible, se tomarán medidas objetivas para la evaluar la obtención de la competencia a través del entrenamiento.
- 2.6** El coordinador mantendrá un registro actualizado del entrenamiento que ha recibido cada miembro del personal. Estos registros proveerán evidencia que los miembros del personal hayan sido entrenados adecuadamente y que su competencia para realizar las tareas asignadas hayan sido evaluadas.
- 2.7** El coordinador será responsable de demostrar que la experiencia y competencia técnica de sus colaboradores son suficientes para las tareas asignadas y que cumplen con los requisitos técnicos señalados en este documento.

### **3. Organización y diseño**

#### **3.1 Programa**

**3.1.1** El coordinador identificará y planeará aquellos procesos que afectan directamente la calidad del ensayo de aptitud y asegurará que ellos se llevan a cabo de acuerdo con procedimientos preestablecidos.

Antes de comenzar el ensayo se acordará y documentará un plan que habitualmente incluirá la siguiente información.

- a) nombre y dirección del coordinador del ensayo de aptitud;
- b) nombre y dirección del coordinador y de otras personas involucradas en el diseño y operación del ensayo de aptitud;

- c) naturaleza y propósito del ensayo de aptitud
- d) procedimiento para seleccionar los participantes del ensayo de aptitud o criterio a ser cumplido antes que la participación sea permitida, si corresponde
- e) nombre y dirección de los colaboradores involucrados en la provisión del ensayo de aptitud (por ejemplo, procesamiento de la muestra, ensayo de homogeneidad, determinación del valor asignado y su incertidumbre);
- f) cantidad e identidad de los participantes esperados en el ensayo de aptitud;
- g) descripción de la forma en que las unidades a ensayar serán obtenidas, procesadas, verificadas y distribuidas;
- h) descripción de la información que será suministrada a los participantes (prenotificación) y el cronograma de las varias fases del ensayo;
- i) la fecha inicial esperada y final del ensayo, incluyendo, cuando sea apropiado, de las fechas en las que el ensayo será llevado a cabo por los participantes;
- j) para los ensayos en marcha, la frecuencia o fechas en las cuales las unidades de ensayo serán distribuidas a los participantes
- k) información sobre los métodos o procedimientos que los participantes utilizarán para realizar el ensayo o mediciones (comúnmente sus ensayos de rutina);
- l) un programa de los métodos estadísticos a ser usados, incluyendo la determinación de valores asignados e incertidumbre y de técnicas para detectar valores atípicos o de disminuir los efectos de los resultados extremos;
- m) descripción de los datos e información que serán suministrados a los participantes;
- n) bases de las técnicas de evaluación del desempeño, si es apropiado;
- o) descripción del alcance en que serán hechos públicos los resultados del ensayo y las conclusiones basadas en los resultados del ensayo.

**3.1.2** Se establecerá un mecanismo para recibir recomendaciones de cómo planear las actividades involucradas en la realización de cada ensayo de aptitud por todas las partes interesadas incluidos los laboratorios que participarán ,

**3.1.3** Las responsabilidades del grupo consultor, bajo la dirección de un coordinador, incluirá, pero no estará limitado, a las consideraciones siguientes;

- a) nominación de los ensayos más significativos a ser aplicados en las unidades de ensayo;
- b) diseño del ensayo;
- c) naturaleza de la unidad de ensayo y ensayo seleccionado, como también una breve descripción de las consideraciones básicas de esta elección, si es apropiado;
- d) método de ensayo a ser empleado, si es apropiado;
- e) dificultades esperadas en la preparación y mantenimiento de la homogeneidad de las unidades a ensayar;
- f) preparación de instrucciones detalladas para los participantes;
- g) preparación de formularios de informes normalizativos a ser usados por los participantes;
- h) número de dígitos significativos a los cuales deben responder los resultados
- i) comentarios sobre cualquier dificultad técnica encontrada por los participantes;
- j) asesoramiento en la evaluación de la competencia técnica de los laboratorios participantes;
- k) medidas apropiadas para juzgar la competencia de los participantes;
- l) comentarios sobre el desempeño, tanto de los participantes individuales como de los participantes en general;
- m) comentarios técnicos sobre el informe del resumen;

- n) evaluación de las respuestas de los participantes con pobre desempeño (si se requiere la retroalimentación)

### **3.2 Preparación de las unidades de ensayo**

**3.2.1** En el planeamiento de todo el proceso para la preparación, ensayo y distribución de los materiales de ensayo y de las unidades de ensayo, el coordinador proporcionará, si es apropiado, procedimientos y recursos para:

- a) selección de material;
- b) mantenimiento de un ambiente apropiado para la preparación y ensayo del material de ensayo;
- c) preparación del material;
- d) medición y ensayo;
- e) evaluación de la homogeneidad del material de ensayo;
- f) evaluación de la estabilidad del material de ensayo;
- g) asegurar las adecuadas instalaciones y condiciones para el almacenamiento;
- h) asegurar el adecuado embalaje y etiquetado;
- i) asegurar el transporte y arreglos de distribución apropiado

**3.2.2** El coordinador será capaz de demostrar que el material de ensayo es suficientemente homogéneo para el ensayo de aptitud particular

**3.2.3** Cuando se producen matrices de materiales de ensayo, estas deberán tener, si es practicable, la misma matriz o casi la misma que los materiales de ensayo de rutina para simular el proceso de medición tan real como sea posible.

### **3.3 Homogeneidad y ensayo de estabilidad**

**3.3.1** Cuando sea apropiado, el coordinador o sus colaboradores usarán una selección estadística aleatoria, de un número de muestras preparadas de una cantidad de material de ensayo, para evaluar la homogeneidad del material. Este procedimiento de evaluación será documentado.

**3.3.2** La evaluación de la homogeneidad será realizada después que el material de ensayo haya sido envasado en su forma final y antes de su distribución a los participantes

**3.3.3** Se demostrará que las muestras del material de ensayo son suficientemente estables para asegurar que no sufrirán ningún cambio significativo durante el ensayo de aptitud.

### **3.4 Diseño estadístico**

**3.4.1** El coordinador documentará el modelo estadístico y las técnicas de análisis de los datos a ser usados, junto con una descripción de las razones para su selección y asegurará que las mismas se realizan de acuerdo con procedimientos prescritos.

**Nota:** Los procedimientos estadísticos recomendados para el tratamiento de los datos de los ensayos de aptitud se establecen en el documento: "Métodos estadísticos utilizados en los ensayos de aptitud mediante comparaciones interlaboratorios"; el cual forma parte de esta metodología general.

**3.4.2** El diseño estadístico apropiado de un ensayo de aptitud es esencial. En el diseño de un ensayo el coordinador dará una cuidadosa consideración a lo siguiente;

- a) la precisión y veracidad requerida o esperada del método de ensayo;
- b) el número de participantes en el ensayo;
- c) el número de muestras a ser ensayadas y el número de réplicas de ensayo o de mediciones a ser realizadas para cada muestra o por cada determinación;
- d) el procedimiento a ser usado para estimar el valor asignado de cada mensurando y su incertidumbre;
- e) procedimiento a ser usado para identificar valores atípicos (aberrantes);
- f) si es apropiado, la homogeneidad y estabilidad de los materiales de ensayo;
- g) procedimiento para determinar la desviación típica para la evaluación de la aptitud

Nota: En ausencia de información confiable concerniente en a), puede ser necesario realizar un ensayo de comparación ínterlaboratorio piloto para obtenerla.

#### **4. Selección del método o procedimiento**

- 4.1** A los participantes del ensayo normalmente se les permitirá usar el método de ensayo o procedimiento de medición de su elección, que sea consistente con los procedimientos de rutina usados en sus laboratorios. En ciertas circunstancias el coordinador del ensayo puede instruir a los participantes para usar un método especificado. Tales métodos son generalmente métodos normalizados: aceptados nacional o internacionalmente y deben haber sido validados por un procedimiento apropiado.
- 4.2** Cuando a los participantes les sea permitido usar un método de su elección, el coordinador, si considera el método apropiado, requerirá detalles del mismo, para permitir la comparación y el comentario de los resultados obtenidos por los diferentes métodos de ensayo.

#### **5. Conducción de un ensayo de aptitud**

##### **5.1 Instrucciones a los participantes (véase también 7.1)**

- 5.1.1** El coordinador dará a los participantes aviso anticipado de la intención de realizar un ensayo de aptitud para asegurar que están enterados de la intención del mismo y que su personal estará disponible en el momento requerido.
- 5.1.2** El coordinador dará instrucciones documentadas a todos los participantes. Tales instrucciones pueden, por ejemplo, ser incluidas como una parte integral del proceso del ensayo.
- 5.1.3** Las instrucciones a los participantes incluirán detalles de los factores que puedan influir en el ensayo de los materiales, por ejemplo, condiciones de almacenamiento, la naturaleza de los materiales o unidades de ensayo, el procedimiento de ensayo empleado y el momento apropiado para la realización del ensayo.
- 5.1.4** Las instrucciones específicas de la forma de registrar es informar los resultados del ensayo incluirán, pero no estarán limitadas a: las unidades de medición, el número de cifras significativas, bases para el informe y la fecha final para recibir los resultados del ensayo.
- 5.1.5** A los participantes del ensayo se les requerirá que apliquen a las muestras del ensayo de aptitud un tratamiento similar al aplicado a las muestras de rutina (a

menos que haya requisitos particulares para el ensayo de aptitud que necesiten apartarse de este principio).

- 5.1.6** Los valores asignados no serán revelados a los participantes hasta después que los resultados hayan sido comparados.

## **5.2 Manipulación y almacenamiento de los materiales**

- 5.2.1** Para evitar la contaminación del material de ensayo, el coordinador y cualquiera de sus colaboradores asociados identificarán, preservarán y segregarán todos los materiales de ensayo y unidades de ensayo, por ejemplo, de todos los productos químicos y otros materiales desde el momento de preparación hasta su distribución a los participantes.

- 5.2.2** El coordinador y cualquier colaborador asociado asegurarán un embalaje adecuado de todos los materiales de ensayo y proveerán áreas de almacenamiento seguro y/o locales de depósitos que prevengan de daños y deterioros a cualquier unidad o material entre la preparación y la distribución. Tales áreas serán definidas para los despachos autorizados y recepción, según los métodos apropiados.

- 5.2.3** Cuando sea apropiado, la condición de todas las unidades y materiales almacenados o depositados, serán evaluados a intervalos específicos, durante su período de almacenamiento, para detectar los posibles deterioros.

## **5.3 Embalaje, marcado y distribución**

- 5.3.1** El coordinador controlará, con la extensión necesaria, los procesos de embalaje y de marcado para asegurar la conformidad con los requisitos relevantes de seguridad y transporte.

Nota 1: Se tomarán los pasos adecuados para asegurar que las muestras presentadas mantienen su integridad, por ejemplo, aquellas que requieren un almacenamiento ininterrumpido en condiciones de frío o que no deben estar expuestas a los rayos X, golpes o vibraciones. La mayor parte de los materiales químicos se beneficiarán con embalajes herméticos, para evitar la contaminación atmosférica, por ejemplo, vapores de combustibles o escapes de motores que puedan presentarse durante el transporte.

Nota 2: En algunos programas de ensayos, tales como comparaciones de mediciones, a los laboratorios participantes se les requiere transportar las unidades hacia otros participantes. En tales casos los laboratorios serán previstos con las instrucciones documentadas para el transporte de las unidades.

- 5.3.2** El coordinador asegurará que las etiquetas de los materiales están adheridas a los envases de los productos de las unidades individuales y están diseñadas para permanecer legibles e intactas dentro del período de uso del ensayo de aptitud.

## **6. Análisis de datos e interpretación de los resultados**

### **6.1 Análisis de datos y registros**

- 6.1.1** Los equipos de procesamiento de datos serán los adecuados para todos los requisitos de las entradas de datos y el análisis estadístico y serán capaces de

proveer resultados oportunos y válidos. El coordinador también establecerá y mantendrá una descripción de todos los equipos de procesamiento de datos.

- 6.1.2 El coordinador definirá el rol y responsabilidad y designará una persona para ser responsable de la operación efectiva del sistema de procesamiento de datos.
- 6.1.3 Todos los equipos de procesamiento de datos y sistemas de software serán mantenidos y validados apropiadamente de acuerdo con los procedimientos documentados antes de ser puestos en uso. Se registrarán los resultados de tales mantenimientos y verificaciones operacionales.
- 6.1.4 Los resultados recibidos de los participantes serán registrados y analizados prontamente mediante procedimientos estadísticos documentados apropiados.
- 6.1.5 El análisis de datos generará resúmenes de mediciones y estadísticos de desempeño e información asociada consistente con el modelo y objetivos del ensayo de aptitud.
- 6.1.6 La influencia de los resultados extremos en el resumen estadístico será minimizada por el uso de décimas apropiadas para detectar valores estadísticos atípicos o por el uso de estadísticos robustos.

El coordinador tendrá criterios y procedimientos documentados para manipular los resultados de ensayo que puedan ser inapropiados para la evaluación estadística, por ejemplo, errores grandes, disparatados, errores de cálculos y transposiciones.

- 6.1.7 El proveedor dispondrá de un criterio documentado para determinar si las unidades de ensayo no son apropiadas para la evaluación, por ejemplo, por falta de homogeneidad, inestabilidad o contaminación indetectable.

## **6.2 Evaluación del desempeño**

- 6.2.1 Cuando se requiera una evaluación de desempeño, el coordinador de ensayos de aptitud será responsable de asegurar que el método de evaluación sea apropiado para el mantener la credibilidad del ensayo. Tal método estará documentado e incluirá una descripción de las bases empleadas para hacer la evaluación.
- 6.2.2 Si es apropiado, el coordinador reclutará la asistencia de consejeros técnicos, que pueden incluir a especialistas en estadística, para brindar comentarios expertos sobre el desempeño de los participantes con respecto a lo siguiente:

- a) desempeño total contra expectativas previas, tomando en cuenta las incertidumbres de las mediciones
- b) variaciones intralaboratorios e ínterlaboratorios y comparaciones con ensayos previos similares o datos precisos publicados;
- c) variaciones entre métodos o procedimientos, si es aplicable
- d) posibles fuentes de errores (con referencia a resultados extremos) y sugerencias para mejorar el desempeño;
- e) otras sugerencias, recomendaciones y comentarios generales;
- f) conclusiones;

## **6.3 Informes de los programas de ensayos de aptitud**

- 6.3.1 El contenido de los informes de los programas de ensayos de aptitud variarán dependiendo del propósito del ensayo en particular, pero cada informe será

claro y explícito e incluirá los datos de la distribución de los resultados de todos los participantes, además de una indicación del desempeño de los participantes individuales

**6.3.2** La siguiente información será incluida normalmente en los informes de los ensayos de aptitud

- a) nombre y dirección del organismo que conduce o coordina el programa;
- b) nombres y cargos de las personas involucradas en el diseño y conducción del programa de ensayo;
- c) fecha de emisión del informe
- d) número de informe e identificación clara del programa;
- e) descripción clara de las unidades o materiales usados, incluyendo (si es apropiado) los detalles de la preparación de la muestra y del ensayo de homogeneidad
- f) códigos de los laboratorios participantes y resultados de los ensayos;
- g) datos y resúmenes estadísticos, incluyendo los valores asignados y el rango de los resultados aceptables;
- h) procedimientos usados para establecer los valores asignables;
- i) detalles de la trazabilidad e incertidumbre de los valores asignados, si es aplicable;
- j) los valores asignados y resumen estadístico de los métodos o procedimientos de ensayo usados por otros participantes (si son usados métodos diferentes por diferentes participantes);
- k) comentarios del coordinador y asesores técnicos sobre el desempeño del participante;
- l) procedimientos para diseñar e implementar el programa (que pueden incluir referencias a protocolos del programa);
- m) procedimientos usados para el análisis estadístico de los datos, si es apropiado;
- n) orientación sobre la interpretación del análisis estadístico cuando sea apropiado.

Nota: Para los ensayos operados regularmente, puede ser suficiente tener informes simplificados, de forma tal que muchos de los elementos recomendados en 6.3.2 pueden ser excluidos de los informes de rutina, pero deben ser incluidos en los protocolos del ensayo o en informes de resúmenes y ser suministrados a pedido de los participantes.

**6.3.3** Los informes estarán disponibles para los participantes en cronogramas especificados. En algunos programas, tales como de comparaciones de mediciones en períodos prolongados, es conveniente emitir informes intermedios a los participantes individuales.

Nota: Sin embargo, idealmente, todos los datos originales suministrados serán informados a todos los participantes, esto es posible que no se logre en algunos ensayos muy grandes. En estos casos los participantes individuales recibirán, al menos, los resultados de todos los participantes, en forma resumida (por ejemplo en forma gráfica)

## **7. Comunicación con los participantes**

**7.1** El coordinador, brindará información detallada a los posibles participantes sobre la participación en el programa de ensayo de aptitud, tal como un protocolo formal del programa. Esto incluirá detalles del alcance del ensayo, aranceles para la participación y la política sobre los laboratorios que pueden participar.

Nota: La comunicación posterior con los participantes puede ser por medio de cartas, boletines y/o informes, además de reuniones abiertas periódicas.

**7.2** El coordinador comunicará prontamente, de forma escrita, a los participantes sobre cualquier cambio en el diseño u operación del programa.

**7.3** Los participantes pueden recurrir al coordinador si discrepan en la evaluación de su desempeño en un ensayo de aptitud

**7.4** La retroalimentación desde los laboratorios debe ser alentada, de forma tal que los participantes contribuyan activamente al desarrollo del programa.

## **8. Confidencialidad**

**8.1** La identidad de los participantes en un programa de ensayo de aptitud será normalmente confidencial (ver nota) y conocida sólo por el mínimo número de personas involucradas en la provisión y evaluación del ensayo.

**8.2** Toda la información suministrada por un participante al coordinador será tratada en forma confidencial

Nota: Los participantes pueden elegir renunciar a la confidencialidad dentro del grupo para propósitos de discusión y asistencia mutua, por ejemplo, para mejorar el desempeño. También se puede renunciar a la confidencialidad por motivos regulatorios o de acreditación. En algunas circunstancias se puede requerir que el organismo coordinador informe los desempeños pobres a una autoridad particular, pero es conveniente que se notifique a los participantes sobre esta posibilidad cuando acuerden participar en el ensayo.

## **9. Colusión y falsificación de resultados**

*Los programas de ensayos de aptitud estarán diseñados para asegurar que la posibilidad de colusión o de falsificación de los resultados sea mínima, si es practicable.*

Nota 1: Sin embargo, el coordinador tomará todas las medidas razonables para prevenir la colusión y será responsabilidad de los participantes el evitarlo.

Nota 2: Se comete colusión entre laboratorios cuando no se envían datos verdaderamente independientes. Los laboratorios también pueden dar una falsa impresión de su desempeño si rutinariamente llevan a cabo análisis únicos, pero informan el promedio de determinaciones repetidas sobre las unidades del ensayo de aptitud o conducen a repeticiones adicionales a aquellas especificadas para un ensayo de aptitud particular.

## **Bibliografía**

- 1. NC-ISO 3534-1. Estadística – Vocabulario y símbolos. Parte 1. Términos de probabilidad y estadística general.**
- 2. NC-ISO/IEC Guía 43-1, Ensayos de aptitud por comparaciones íter laboratorios- Parte 1: Desarrollo y funcionamiento de programas de ensayos de aptitud.**

- 3. NC-ISO/IEC, Guía 43-2, Ensayos de aptitud por comparaciones ínter laboratorios-Parte 2: Selección y uso de programas de ensayos de aptitud por organismos de acreditación de laboratorios.**
- 4. ISO/DIS 13528:2000. Métodos estadísticos utilizados en los ensayos de aptitud mediante comparaciones ínter laboratorios**
- 5. ILAC-G13:2000 Guía sobre los requisitos para la competencia de los proveedores de programas de ensayos de aptitud.**

**Anexo B**  
**Métodos estadísticos utilizados en los ensayos de aptitud mediante comparaciones interlaboratorios**

**Contenido**

1.	Objetivo y alcance.....	12
2.	Referencias normativas .....	12
3.	Definiciones.....	12
4.	Lineamientos estadísticos para el diseño y la interpretación de los ensayos de aptitud. ....	14
4.1	Señales de acción y de advertencia .....	14
4.2	Lineamiento para limitar la incertidumbre del valor asignado. ....	14
4.3	Lineamiento para seleccionar el número de réplicas (mediciones repetidas).....	15
4.4	Homogeneidad y estabilidad de las muestras .....	16
4.5	Métodos de medición definidos operativamente.....	16
4.6	Informe de los datos.....	16
4.7	Período de validez de los resultados de los ensayos de aptitud .....	16
5.	Determinación del valor asignado y su incertidumbre típica.....	17
5.1	Selección del método para determinar el valor asignado. ....	17
5.2	Valores conocidos (Formulación).....	17
5.3	Valores de referencia certificados.....	17
5.4	Valores de referencia .....	18
5.5	Valores de consenso de los laboratorios expertos .....	20
5.6	Valores de consenso de los laboratorios participantes.....	20
5.7	Comparación del valor asignado con el promedio robusto.....	25
5.8	Valores ausentes .....	25
6.	Determinación de la desviación típica para la evaluación de la aptitud.....	26
6.1	A partir de los resultados de un experimento de precisión .....	26
6.2	A partir de los datos obtenidos en una ronda de un ensayo de aptitud .....	26
7.	Cálculo de los estadísticos de desempeño.....	27
7.1	Estimados del sesgo del laboratorio .....	27
7.2	Diferencia porcentual. ....	29
7.3	Clasificaciones y clasificaciones del porcentaje.....	31
7.4	Valores Z.....	32
7.5	Números $E_n$ .....	35
7.6	Valores $Z'$ .....	36
8.	Métodos gráficos para combinar los valores de desempeño tras varias rondas de un programa de aptitud .....	36
8.1	Aplicaciones .....	36
8.2	Gráfico de control de Shewhart.....	37
9.	Programa de ensayo de muestras divididas .....	39
10.	Programas de ensayo con valores conocidos .....	43
11.	Programa de comparación de mediciones.....	43
12.	Pasos a seguir para la aplicación de los métodos estadísticos en los programas de ensayos interlaboratorio .....	44
	ANEXO A_ Símbolos .....	46
	ANEXO B_ Comprobación de la homogeneidad y la estabilidad entre las muestras.....	47
	ANEXO C_ Análisis robusto .....	51
	ANEXO D_ Dócima de Grubbs y de Cochran para determinar valores atípicos .....	53

## **Métodos estadísticos utilizados en los ensayos de aptitud mediante comparaciones interlaboratorio.**

### **1. Objetivo y alcance**

**El objetivo de este trabajo es complementar la Guía 43 de la NC-ISO/IEC mediante:**

- la presentación de descripciones detalladas de métodos estadísticos que se utilizan en el análisis de los datos obtenidos de los programas (esquemas) de ensayos de aptitud;
- la presentación de recomendaciones sobre su utilización en la práctica.

Estos métodos estadísticos nos permiten determinar el desempeño de laboratorios individuales en la realización de ensayos o mediciones específicas y monitorear el desempeño continuo de los laboratorios.

### **2. Referencias normativas**

NC-ISO 3534-1. Estadística- Vocabulario y Símbolos. Parte 1. Términos de probabilidad y estadística general.

ISO 5725-2. Exactitud (veracidad y precisión) de los métodos y resultados de medición. Parte 2: Método básico para determinar la repetibilidad y reproducibilidad de un método de medición normalizado.

NC-ISO/IEC, Guía 43-1, Ensayos de aptitud por comparaciones interlaboratorios- Parte 1. Desarrollo y funcionamiento de programas de ensayos de aptitud.

NC- ISO/IEC, Guía 43-2, Ensayos de aptitud por comparaciones interlaboratorios- Parte 2. Selección y uso de programas de ensayos de aptitud por organismos de acreditación de laboratorios.

ISO/DIS 13528:2000 Métodos estadísticos utilizados en los ensayos de aptitud mediante comparaciones interlaboratorio.

### **3. Definiciones**

Se aplican las definiciones de la norma NC-ISO 3534-1, además de las siguientes tomadas de otras normas.

#### **3.1 Ensayo**

Operación técnica que consiste en determinar una o más característica de un producto, proceso o servicio según un procedimiento específico.

#### **3.2 Unidad de ensayo**

Material o dispositivo presentado al laboratorio participante a los fines del ensayo de aptitud.

#### **3.3 Método de ensayo**

Procedimiento técnico especificado para realizar un ensayo

#### **3.4 Resultado de ensayo**

El valor de una característica obtenido aplicando completamente un método específico de medición.

### **3.5 Comparaciones interlaboratorio**

Organización, realización y evaluación de ensayos o sobre la misma unidad de ensayo o sobre otras similares, por dos o más laboratorios, de acuerdo con condiciones predeterminadas.

### **3.6 Ensayo de aptitud (de un laboratorio)**

Determinación del desempeño de un laboratorio de ensayo mediante comparaciones interlaboratorios

### **3.7 Valor asignado**

Valor atribuido a una magnitud particular y aceptada, a veces por convención, como portador de una incertidumbre apropiada para un fin determinado.

### **3.8 Desviación típica para la evaluación de la aptitud**

Desviación típica utilizada en la evaluación de la aptitud, puede estar relacionada con la desviación típica de reproducibilidad o con la de declaración de la idoneidad del método de medición.

### **3.9 Valor Z**

Medida normalizada del sesgo del laboratorio, que se calcula utilizando el valor asignado y la desviación típica para la evaluación de la aptitud.

### **3.10 Coordinador**

Organismo (o persona) con responsabilidad para coordinar todas las actividades relacionadas con la operación de un programa de ensayo de aptitud

### **3.11 Veracidad**

Proximidad de la concordancia entre el valor promedio, obtenido de una serie larga de resultados de ensayos y un valor de referencia aceptado.

### **3.12 Precisión**

Proximidad de la concordancia entre resultados de ensayos independientes, obtenidos bajo condiciones determinadas (estipuladas). La repetibilidad y la reproducibilidad son medidas de precisión.

### **3.13 Valor atípico**

Miembro de un conjunto de valores, el cual es inconsistente con los otros miembros de dicho conjunto.

### **3.14 Resultados extremos**

Valores atípicos y otros valores los cuales son exageradamente inconsistentes con otros miembros del conjunto de datos.

Nota: Estos resultados pueden tener una influencia profunda sobre los estadísticos finales, tales como la media y la desviación típica.

### **3.15 Técnicas estadísticas robustas**

Técnicas para minimizar la influencia que pueden tener los resultados extremos en las estimaciones de la media y de la desviación típica.

Nota: Estas técnicas asignan menos peso a los resultados extremos, en lugar de eliminarlos de un conjunto de datos.

### 3.16 Incertidumbre de la medición

Parámetro asociado a los resultados de una medición que caracteriza la dispersión de los valores que podrían ser atribuidos al mensurando.

## 4. Lineamientos estadísticos para el diseño y la interpretación de los ensayos de aptitud.

### 4.1 Señales de acción y de advertencia

Se describe algunos sencillos criterios numéricos y gráficos que se deben aplicar a los datos obtenidos en un ensayo de aptitud, para ver si causan señales de acción o advertencia y evaluar el desempeño de los laboratorios individuales y colectivo. Incluso en los laboratorios bien operados, con personal experimentado, se pueden obtener resultados anómalos. Por esta razón, los criterios que aquí aparecen no se deben utilizar para acusar a los laboratorios de ineptos para aplicar el método de medición que se examina.

Los criterios que aquí aparecen están elaborados de modo tal que cuando la desviación típica para la evaluación de la aptitud se basa en el desempeño observado (utilizando uno de los métodos que se describen en 6,1 y 6,2) los criterios emiten señales de acción si los resultados son tan excepcionales que requieren de investigación y acciones correctivas (los coeficientes utilizados en los criterios son apropiados cuando los datos están distribuidos normalmente). Es preciso tomar precauciones cuando los criterios se aplican a datos extremadamente no normales.

En los programas donde la desviación típica para la evaluación de la aptitud se basa en el desempeño observado, cuando un resultado emite una señal de acción el laboratorio decidirá que investigaciones y acciones correctivas resultan apropiadas en consulta con el coordinador o si es necesario, con el organismo de acreditación. Examinará su procedimiento e identificará una o más acciones correctivas que, según el personal del laboratorio puedan evitar que se repitan dichos resultados. El laboratorio le puede pedir consejos al coordinador en cuanto a las posibles causas de su problema, o puede pedirle al coordinador que consulte con otros expertos. El laboratorio tomará parte en las futuras rondas del programa de ensayo de aptitud para evaluar la eficiencia de las acciones correctivas. Entre las acciones correctivas adecuadas pueden encontrarse las siguientes:

- a) comprobar que el personal comprende y aplique el procedimiento de medición;
- b) comprobar que sean correctos todos los detalles del procedimiento de medición;
- c) comprobar la calibración de los equipos y la composición de los reactivos;
- d) sustituir los equipos o reactivos sospechosos;
- e) pruebas comparativas del personal, los equipos y/o los reactivos con los de otros laboratorios.

### 4.2 Lineamiento para limitar la incertidumbre del valor asignado.

El valor asignado  $X$  tiene una incertidumbre típica  $u_X$  que depende del método empleado para obtenerlo, cuando se obtiene a partir de ensayos en varios laboratorios (véase 5.6) del número de estos, y quizás de otros factores. En el apartado 5 aparecen métodos para calcular el valor asignado y su incertidumbre típica.

La desviación típica para la evaluación de la aptitud  $\hat{\sigma}$  se utiliza para evaluar el tamaño de los estimados de sesgos de laboratorio detectados en el ensayo de aptitud. En el apartado 6 aparecen los métodos para obtener la desviación típica para la

evaluación de la aptitud, y en el apartado 7 se ofrecen criterios que la comparan con los estimados del sesgo del laboratorio.

Si la incertidumbre típica  $u_X$  del valor asignado es demasiado grande en comparación con la desviación para el ensayo de aptitud, entonces existe el riesgo de que algunos laboratorios reciban señales de acción y advertencia debido a la inexactitud en la determinación del valor asignado y no debido a algunas causas atribuibles a los laboratorios. Por esta razón, la incertidumbre típica del valor asignado se establecerá y se informará a los laboratorios que participan en programas de ensayos de aptitud.

Si  $u_X \leq 0,3 \hat{\sigma}$

entonces se puede despreciar la incertidumbre del valor asignado y no es necesario incluirla en la interpretación de los resultados del ensayo de aptitud.

Si este lineamiento no se cumple, entonces el coordinador tendrá en cuenta los siguientes pasos:

- a) buscar un método para definir el valor asignado cuya incertidumbre cumpla el lineamiento anterior;
- b) utilizar la incertidumbre del valor asignado en la interpretación de los resultados del ensayo de aptitud, (véase 7.5 en los números  $E_n$  o el 7.6 en el valor  $Z'$ );
- c) informar a los participantes del ensayo de aptitud que no se puede despreciar la incertidumbre del valor asignado.

Como ejemplo sencillo, supongamos que el valor asignado  $X$  se determina como el promedio  $\bar{x}$  de los resultados de ensayos finales en 11 laboratorios, y que la desviación típica para la evaluación de la aptitud se determina como la desviación típica  $s$  de esos mismos 11 resultados de modo que  $\hat{\sigma} = s$ .

La incertidumbre típica del valor asignado en esta situación se puede estimar mediante  $u_X = s / \sqrt{11} = 0,30 s$  de modo que parece haberse cumplido el requisito. Sin embargo, el requisito no se puede cumplir en esta situación con menos de 11 laboratorios. Además, la incertidumbre del valor asignado será mayor que  $s / \sqrt{11}$  si hay sesgo en el método de medición.

Como se evidencia al emplear  $s$  el número mínimo de laboratorios participantes ( $p$ ) se obtiene a partir de  $s / \sqrt{p} = 0,3 s$  donde  $p = (1 / 0,3)^2 = 3,33^2 \approx 11$

#### **4.3 Lineamiento para seleccionar el número de réplicas (mediciones repetidas)**

La repetibilidad contribuye a la variación entre los sesgos de laboratorio en un ensayo de aptitud. Si la repetibilidad es demasiado grande en comparación con la desviación típica para el ensayo de aptitud, entonces existe riesgo de dicha variación ocasione errores en el ensayo de aptitud. Por lo que puede suceder que algunos laboratorios tendrán grandes sesgos de laboratorio en una ronda, pero no en la siguiente, y tendrán dificultades para identificar la causa.

Por esta razón si resulta conveniente limitar la influencia de la repetibilidad, el número de réplicas  $n$  realizadas por cada laboratorio en un ensayo de aptitud será tal que:

$$\sigma_r / \sqrt{n} \leq 0,3 \hat{\sigma}$$

Donde  $\sigma_r$  es la desviación típica de la repetibilidad que se ha establecido en un experimento interlaboratorio anterior o en el caso que no se posea, se estima ( $s_r$ ) a partir de los datos del ensayo de aptitud. (Véase 5.6.2).

Si no se cumple el requisito anterior, entonces se aumentará el número de réplicas para las rondas posteriores de modo que el mismo se cumpla. Este enfoque asume que los laboratorios tienen por lo general una repetibilidad similar.

#### **4.4 Homogeneidad y estabilidad de las muestras**

En el Anexo B aparecen métodos para comprobar si las muestras que se utilizarán en un ensayo de aptitud son adecuadas, homogéneas y estables. Cuando no se cumple el criterio de homogeneidad del Anexo B, la desviación típica para la evaluación de la aptitud incluirá un margen para la heterogeneidad de las muestras, según se describe en el Anexo B.

#### **4.5 Métodos de medición definidos operativamente**

Los participantes son libres de escoger entre métodos definidos operativamente cuando no lleguen a un consenso válido entre ellos. Existen dos formas de evitar este problema.

- a) se utiliza un método normalizativo para establecer el valor asignado, y se orienta a los participantes que apliquen el mismo método;
- b) se produce un valor independiente del valor asignado para cada método, utilizar esta variante no es la más recomendable por ser mucho más complejo y difícil de aplicar.

#### **4.6 Informe de los datos**

Teniendo en cuenta los cálculos requeridos para el ensayo de aptitud, se recomienda no redondear los resultados de medición individuales en más de  $\sigma_r / 2$

Los resultados de la medición no serán truncados, se le pedirá a los participantes que informen los valores actuales de sus resultados individuales de medición, y también se les informará que si informan resultados truncados sobre la muestra, entonces todos sus datos para esa muestra se excluirán del análisis.

#### **4.7 Período de validez de los resultados de los ensayos de aptitud**

El período de validez del resultado obtenido por un laboratorio es una ronda única del programa de ensayos de aptitud se limita al momento que tomó el laboratorio para realizar el ensayo. De este modo, si el laboratorio alcanza un resultado satisfactorio en una ronda única, dicho resultado no se utilizará para afirmar que el laboratorio obtuvo datos confiables en alguna otra ocasión.

El laboratorio que opera un sistema de calidad y tiene antecedentes de resultados satisfactorios en muchas rondas en un programa de ensayo de aptitud estará facultado para utilizar dichos resultados como evidencia de que es capaz de obtener datos consistentemente confiables.

## **5. Determinación del valor asignado y su incertidumbre típica.**

### **5.1 Selección del método para determinar el valor asignado.**

En los apartados del 5.2 al 5.6 se describen cinco formas de determinar el valor asignado X. La selección de uno de estos métodos será responsabilidad del coordinador en consulta con los miembros del programa.

La determinación del valor asignado será responsabilidad del coordinador. El valor asignado no se revelará a los participantes hasta que hayan informado sus resultados al coordinador, el cual preparará un informe con los detalles de cómo se obtuvo el valor asignado, las identidades de los laboratorios que participan en su determinación, y (si es posible) declaraciones sobre la incertidumbre típica de la medición del valor asignado.

### **5.2 Valores conocidos (formulación)**

Con resultados determinados por una formulación específica de la unidad de ensayo (por ejemplo fabricación o dilución)

El material de ensayo se puede preparar mezclando los constituyentes en proporciones específicas, o agregando una proporción especificada de un constituyente en un material básico. En este caso, el valor asignado X se obtiene calculando a partir de las masas utilizadas.

Este enfoque es especialmente valioso si se pueden preparar muestras individuales de esta manera y si lo que se va a determinar es la proporción de los constituyentes o del agregado, pues entonces no es necesario preparar una gran cantidad de material y asegurar que es homogénea. No obstante, cuando las formulaciones ofrecen muestras en las que la adición no tiene afinidad tan fuerte como en los materiales típicos o los presenta de forma diferente, puede ser más conveniente utilizar otro enfoque.

La limitación de este método (en el análisis químico) es que resulta necesario tomar precauciones para garantizar que:

- a) el material básico está realmente libre del constituyente agregado, o que se conoce con exactitud la proporción del constituyente agregado al material básico;
- b) los constituyentes están mezclados de forma homogénea (cuando se requiere);
- c) están identificadas todas las fuentes de error;
- d) no existe interacción entre los constituyente y la matriz.

#### **Ejemplo:**

#### **Determinación del contenido de cemento del hormigón endurecido:**

En este caso, se pueden preparar especímenes de hormigón pesando cantidades de los constituyentes (cemento, agregados y agua) y mezclándolos para formar las muestras de hormigón. El enfoque resulta satisfactorio porque la veracidad con la que se pueden preparar los especímenes es mucho mayor que la obtenida con el método analítico utilizado para determinar el contenido de cemento.

### **5.3 Valores de referencia certificados**

Cuando el material que se utiliza en un ensayo de aptitud es un material de referencia certificado (MRC), su valor de referencia certificado se utiliza como el valor asignado X.

#### **Incertidumbre típica $u_X$ del valor asignado:**

Si se utiliza un MRC como material de ensayo, la incertidumbre típica del valor asignado se obtiene de la información sobre la incertidumbre que ofrece el certificado. La limitación de este enfoque es que puede resultar costoso entregar a cada participante del ensayo de aptitud una muestra de un MRC.

#### 5.4 Valores de referencia

En este enfoque, se preparan primero las muestras del material de ensayo que se utilizará como material de referencia (MR) y quedan listo para la distribución entre los participantes. Luego se seleccionan aleatoriamente un número de muestras (10 ó más, lo ideal sería 20) y se ensayan en un laboratorio junto con las muestras del MRC, utilizando un método de medición adecuado y en condiciones de repetibilidad. Entonces el valor asignado  $X_{MR}$  del material de ensayo se calcula a partir de una calibración con respecto a los valores de las muestras del material de referencia certificado (MRC) de la forma que se señala en el ejemplo.

##### **Incertidumbre típica $u_X$ del valor asignado:**

La incertidumbre típica del valor asignado se calcula a partir de los resultados de ensayo realizados en ese material y a la MRC, de valore similares, véase el ejemplo. Este método permite establecer el valor asignado que sea trazable al valor certificado de la MRC, con una incertidumbre típica que se puede calcular y evita el costo de distribuir las muestras del MRC a todos los participantes. Estas son buenas razones para preferirlo sobre otros métodos.

##### **Ejemplo:**

El siguiente ejemplo muestra cómo el valor asignado de un material de ensayo se establece mediante comparaciones directas con un MRC único así como la incertidumbre requerida. Se tiene un MR y MRC que son materiales similares ( en matriz, composición y nivel de resultados). Esta determinación requiere la realización de una serie de ensayos en un laboratorio con muestras de los dos materiales, como se señaló anteriormente, utilizando el mismo método de medición y en condiciones de repetibilidad.

Se registran y se procesan los datos de acuerdo a la tabla 1, determinando las diferencias promedios de los resultados de ensayo de las muestras de MR y MRC, la diferencia promedio ( $\bar{D}$ ), la desviación típica de las diferencias promedios y la incertidumbre típica de las mismas.

La incertidumbre típica de las diferencias promedio ( $u_D$ ) se calcula mediante:

$$u_D = s_D / \sqrt{p}$$

donde:

$s_D$  desviación típica de las diferencias promedio

$p$  número de ensayos en paralelo de la serie de ensayos

Posteriormente se calcula:

$$X_{MR} = X_{MRC} + \bar{D}$$

donde:

$X_{MRC}$  valor asignado del material de referencia certificado

$X_{MR}$  valor asignado del material de referencia

$\overline{D}$  diferencia promedio entre los dos materiales, obtenidos de la serie de ensayos

La incertidumbre típica del valor asignado de MR se calcula:

$$u_{X:MR} = \sqrt{u_{X:MRC}^2 + u_D^2}$$

con los resultados de la tabla 1 y considerando que  $X_{MRC} = 21,62$  y  $u_{X:MRC} = 0,26$

$$X_{MR} = 21,62 + 1,73 = 23,35 \text{ y } u_{X:MR} = \sqrt{0,26^2 + 0,24^2} = 0,35$$

**Tabla 1. Cálculo de la diferencia promedio entre un MRC y un MR, y de la incertidumbre típica de esta diferencia.**

Ensayo en Paralelo	Material de Referencia		Material de Referencia Certificado		Diferencia Promedio
	Ensayo 1	Ensayo 2	Ensayo 1	Ensayo 2	MR-MRC
1	20,5	20,5	19,0	18,0	
2	21,1	20,7	19,8	19,9	2,00
3	21,5	21,5	21,0	21,0	1,05
4	22,3	21,7	21,0	20,8	0,50
5	22,7	22,3	20,5	21,0	1,10
6	23,6	22,4	20,3	20,3	1,75
7	20,9	21,2	21,5	21,8	2,70
8	21,4	21,5	21,9	21,7	-0,60
9	23,5	23,5	21,0	21,0	-0,35
10	22,3	22,9	22,0	21,3	2,50
11	23,5	24,1	20,8	20,6	0,95
12	22,5	23,5	21,0	22,0	3,10
13	22,5	23,5	21,0	21,0	1,50
14	23,4	22,7	22,0	22,0	2,00
15	24,0	24,2	22,1	21,5	1,05
16	24,5	24,4	22,3	22,5	2,30
17	24,8	24,7	22,0	21,9	2,05
18	24,7	25,1	21,9	21,9	2,80
19	24,9	24,4	22,4	22,6	3,00
20	27,2	27,0	24,5	23,7	2,15
					3,00
Diferencia promedio					1,73
Desviación típica					1,07
Incertidumbre típica $(s/\sqrt{20}) - u_D$					0,24

## 5.5 Valores de consenso de los laboratorios expertos

Es conveniente que los laboratorios expertos tengan competencia demostrable en la determinación del (de los) mensurando (s) bajo ensayo, usando métodos validados que se sabe que son altamente precisos y exactos y comparables a los métodos de uso general. Los laboratorios pueden ser en algunos casos, laboratorios de referencia. Al igual que con el enfoque “valores de referencia” (5.4), primero se preparan muestras del material de ensayo lista para ser distribuidas a los participantes. Luego se escogen aleatoriamente algunas de estas muestras para ser analizadas por un grupo de laboratorios expertos. De forma alterna, el grupo de laboratorios expertos puede participar en una ronda de un programa de ensayos de aptitud, y el valor asignado y su incertidumbre se obtendrá una vez terminada la ronda. El número de laboratorios expertos debe ser 10 ó mayor de 10, cantidad menor de laboratorios resultaría que el método pierde eficiencia.

El valor asignado  $X$  se calcula como el promedio robusto de los resultados reportados por el grupo de laboratorios especializados que se obtienen a partir del Algoritmo A del Anexo C. También es posible utilizar las dójimas de Grubbs y de Cochran para eliminar los valores atípicos, siendo la media y la desviación típica de los valores no atípicos el valor asignado y la desviación típica respectivamente. Lo más adecuado es utilizar el Algoritmo A.

### **Incertidumbre típica $u_X$ del valor asignado:**

La incertidumbre típica del valor asignado se estimará como se describe en 5.6.

## 5.6 Valores de consenso de los laboratorios participantes.

El número de laboratorios participantes serán  $\geq 15$ . En los cálculos se toman en cuenta los efectos de los valores extremos utilizando la media robusta y la desviación típica robusta o eliminando los valores atípicos.

### **5.6.1 Teniendo en cuenta los efectos de los valores extremos**

En este enfoque el valor asignado  $X$  para el material de ensayo utilizado en una ronda de un programa de ensayo de aptitud es el promedio robusto de los resultados reportados por todos los participantes en la ronda, calculado a partir del Algoritmo A del Anexo C.

Este enfoque puede resultar particularmente útil con un método de medición empírico, siempre que dicho método esté normalizado

### **Incertidumbre típica $u_X$ del valor asignado:**

Si el valor asignado se calcula como un promedio robusto a partir del Algoritmo A, la incertidumbre típica del valor asignado  $X$  se estima como:

$$u_X = 1,23 \times s^* / \sqrt{p}$$

Donde  $s^*$  es la desviación típica robusta de los resultados calculados usando Algoritmo A en Anexo C. (En este caso, un resultado para un participante es el promedio de todas sus mediciones (réplicas) sobre el material de ensayo)

Las limitaciones de este enfoque son:

- a) puede que no se alcance un consenso real entre los participantes;
- b) el consenso puede presentar sesgo debido al uso general de una metodología con fallas, y este sesgo no se reflejará en la incertidumbre típica del valor asignado calculado según se describió anteriormente.

### **Ejemplo:**

En la tabla 2 aparecen datos de una ronda de un ensayo de aptitud en el que se determinaron las concentraciones de un anticuerpo  $I_g E$  ( inmunoglobulina E ), se realizaron dos ensayos (réplicas) para obtener el resultado de ensayo.

Para aplicar el Algoritmo A, primero se clasificaron los datos en orden ascendente, se calculan los estimados robustos iniciales del promedio y luego se aplica el método iterativo del Algoritmo. La tabla 3 muestra estos cálculos para los resultados del anticuerpo  $I_g E$  que aparecen en la tabla 2.

Los cálculos requeridos por el Algoritmo A se pueden realizar de la forma siguiente:

**Paso 1:** Introducir los datos en una columna en orden ascendente, como se muestran en la Tabla 3 para la Repetición 0. Calcular su promedio y su desviación típica (10, 91 y 3,13 en la tabla 3). Calcular los valores iniciales para el promedio robusto y la desviación típica robusta (10,85 y 3,53 en la tabla 3) utilizando las fórmulas que aparecen el apartado C.1 del Anexo C.

**Paso 2:** Copiar los datos en la siguiente columna según se muestra en la Tabla 3 para la Repetición 1. Utilizar los valores iniciales del promedio robusto y la desviación típica robusta para calcular los valores de corte (5,56 y 16,15 en la tabla 3) por medio de las fórmulas que aparecen en el apartado C.1. Sustituir los datos fuera de los valores de corte por los valores de corte (2,18 se sustituye por 5,56 y 16,30 se sustituye por 16,15). Calcular el nuevo promedio y la nueva desviación típica de los datos alterados (11,03 y 2.81 en la tabla 3). De acuerdo con las fórmulas del Anexo C, el promedio robusto es igual a este promedio (11,03) y la desviación típica robusta (3,19) se obtiene multiplicando la desviación típica por 1,134.

**Paso 3.** No es necesario crear otras columnas de datos. En vez de eso, cambiar los cálculos de los valores de corte en la parte superior de la columna de datos de modo que utilicen el promedio robusto y la desviación típica robusta de la parte inferior de esa misma columna. Esto nos dará los valores de corte (6,24 y 15,82) que aparecen en la Tabla 3 bajo la repetición 2. Entonces se pueden completar los cálculos si se siguen sustituyendo los datos fuera de los valores límites por los valores límites hasta que haya convergencia en las repeticiones. Si se sustituyen los datos, se actualizará automáticamente los promedios y las desviaciones típicas y los valores de corte, pero los cambios en estos valores se reducirán progresivamente hasta que dejen de ser significativos.

Tabla 2 Concentraciones de un anticuerpo I<sub>g</sub> E

Laboratorio	Concentración (x)
A	11,30
B	8,29
C	11,90
D	15,60
E	13,40
F	12,50
G	10,40
H	9,38
I	14,20
J	12,10
K	8,10
L	10,80
M	13,80
N	7,0
O	10,85
P	2,18
Q	8,39
R	6,95
S	11,80
T	10,90
U	16,30
V	9,71
W	10,50
X	13,60
Y	10,10
Z	16,07
a	8,47
p	27
Promedio robusto $x^*(X)$	11,03
Desviación típica robusta $s^*(\hat{\sigma})$	3,04
Nota: El valor $x^*$ y $s^*$ se determina de acuerdo con la tabla 3	

Tabla 3. Cálculos del promedio robusto y la desviación típica robusta para un anticuerpo I<sub>g</sub> E

Repetición	0	1	2	3	4	5
$\delta = 1,5 s^*$	-	5,30	4,79	4,62	4,58	4,56
$x^* - \delta$	-	5,56	6,24	6,41	6,45	6,47
$x^* + \delta$	-	16,15	15,82	15,65	15,61	15,59
P	2,18	5,56	6,24	6,41	6,45	6,47
R	6,95	6,95	6,95	6,95	6,95	6,95
N	7,00	7,00	7,00	7,00	7,00	7,00
K	8,10	8,10	8,10	8,10	8,10	8,10
B	8,29	8,29	8,29	8,29	8,29	8,29
Q	8,39	8,39	8,39	8,39	8,39	8,39
a	8,47	8,47	8,47	8,47	8,47	8,47
H	9,38	9,38	9,38	9,38	9,38	9,38
V	9,71	9,71	9,71	9,71	9,71	9,71
Y	10,10	10,10	10,10	10,10	10,10	10,10
G	10,40	10,40	10,40	10,40	10,40	10,40
W	10,50	10,50	10,50	10,50	10,50	10,50
L	10,80	10,80	10,80	10,80	10,80	10,80
O	10,85	10,85	10,85	10,85	10,85	10,85
T	10,90	10,90	10,90	10,90	10,90	10,90
A	11,30	11,30	11,30	11,30	11,30	11,30
S	11,80	11,80	11,80	11,80	11,80	11,80
C	11,90	11,90	11,90	11,90	11,90	11,90
J	12,10	12,10	12,10	12,10	12,10	12,10
F	12,50	12,50	12,50	12,50	12,50	12,50
E	13,40	13,40	13,40	13,40	13,40	13,40
X	13,60	13,60	13,60	13,60	13,60	13,60
M	13,80	13,80	13,80	13,80	13,80	13,80
I	14,20	14,20	14,20	14,20	14,20	14,20
D	15,60	15,60	15,60	15,60	15,60	15,59
Z	16,07	16,07	15,82	15,65	15,61	15,59
U	16,30	16,15	15,82	15,65	15,61	15,59
Promedio	10,91	11,03	11,03	11,03	11,03	11,03
Desv. Típica	3,13	2,81	2,72	2,69	2,68	2,68
Nuevo $x^*$	10,85	11,03	11,03	11,03	11,03	11,03
Nuevo $s^*$	3,53	3,19	3,08	3,05	3,04	3,04

## 5.6.2 Eliminando los valores atípicos

En este procedimiento se aplica a los datos las dcimas Cochran y de Grubbs. La primera columna contendr el valor del primer ensayo y la segunda columna el valor del segundo ensayo de la muestra de los laboratorios que participan en el ensayo de aptitud, en la tercera columna, los resultados de ensayo y en la cuarta columna, los recorridos de las rplicas. Posteriormente se aplica la dcima de Cochran a los recorridos y la dcima de Grubbs a los resultados. Se eliminan los datos con valores atpicos y se calculan el promedio de los laboratorios  $\bar{x}$  y la desviacin tpica  $s$  de los resultados restantes. En este caso el valor asignado es el promedio calculado y la incertidumbre tpica:

$$u_x = s / \sqrt{p}$$

donde  $s$  es la desviacin tpica de los resultados restantes.

En la tabla 4 se da un modelo de registro de los datos. El procedimiento para la aplicacin de las dcimas de Cochran y de Grubbs se dan en el Anexo D.

En este trabajo al aplicar la dcima de Grubbs los resultados  $\bar{x}_i$  se debe tener en cuenta que el nmero  $n$  de resultados corresponde al nmero  $p$  de laboratorios, por consiguiente en la tabla 17 debe considerarse  $n$  como  $p$  y no como el nmero de rplicas.

Tabla 4. Modelo para 2 ensayos por laboratorio

Laboratorio	Ensayo 1	Ensayo 2	Resultado	Recorrido
Cdigo de los laboratorios	$x_1$	$x_2$	$\bar{x}_1$	$R_1$
	$x_1$	$x_2$	$\bar{x}_2$	$R_2$
	$x_1$	$x_2$	$\bar{x}_3$	$R_3$
	.	.	.	.
	.	.	.	.
	$x_1$	$x_2$	$\bar{x}_p$	$R_p$
			$\bar{x}$	$\bar{R}$

*Este mtodo es menos eficiente que el procedimiento sealado en el apartado 5.6.1, dando lugar a que los errores en los clculos puedan ser mayores, pero las determinaciones de los valores del valor asignado y la incertidumbre tpica resulta ms fcil.*

Si se desea tener una estimacin de la desviacin tpica de la repetibilidad a partir de este mtodo se emplea la expresin:

$$s_r = \sqrt{\sum R_i^2 / 2 p}$$

Para estimar la desviación típica de la reproducibilidad se emplea la expresión:

$$s_R = \sqrt{s_L^2 + s_r^2}$$

En este caso  $s_L = \sqrt{\hat{\sigma}^2 - s_r^2/2}$  y  $\hat{\sigma}^2 = s^2$

El procedimiento señalado anteriormente se aplica cuando  $n=2$ , en el caso que  $n \geq 2$  es más adecuado emplear las desviaciones típicas en vez de los recorridos, por tanto, deberán realizarse las siguientes modificaciones: en la tabla 4 se sustituyen en la última columna los recorridos por las varianzas  $S^2$ , se aplica la dócima de Cochran a las varianzas y la repetibilidad se calcula por la expresión:

$$s_R = \frac{\sum s_i^2}{p}$$

### 5.7 Comparación del valor asignado con el promedio robusto.

Si se aplica el método descrito en los apartados 5.2 al 5.6 para establecer el valor asignado  $X$  tras cada ronda de un programa de ensayos de aptitud, el promedio robusto  $x^*$  de los resultados de la ronda se comparará con el valor asignado. La incertidumbre de la diferencia  $x^* - X$  se estimará como:

$$\sqrt{\frac{s^{*2}}{p}} \times u_X^2$$

donde  $s^*$  es la desviación típica robusta y  $p$  es el número de laboratorios. Si la diferencia es más del doble de su incertidumbre, se buscará la causa. Las posibles causas son:

- sesgo en el método de medición;
- un sesgo común en los resultados de los laboratorios;
- no apreciación de las limitaciones del método si se utiliza el enfoque de valores conocidos;
- sesgo en los resultados.

### 5.8 Valores ausentes

Todos los participantes deben obtener el mismo número de mediciones repetidas (réplicas), pero puede suceder que un laboratorio tenga menos mediciones debido por ejemplo, si se malogran algunos ensayos y no se puede repetir. Si esto sucede, se recomienda el siguiente procedimiento:

Si un laboratorio reporta al menos  $0,59n$  mediciones repetidas, entonces el promedio y la desviación típica de las mediciones se incluirán en los cálculos y se tratarán como si hubieran reportado  $n$  mediciones. El informe mencionaría cuantas mediciones reportaron.

Si el laboratorio reporta menos de  $0,59n$  mediciones repetidas, entonces sus resultados no se incluirán en el cálculo de los estadísticos que influyan en otros laboratorios. Por ejemplo, sus resultados no se incluirán en el cálculo del valor asignado, ni en el cálculo de la desviación típica para la evaluación según se describe en el apartado 5.5 y 6.2 respectivamente. Sus resultados se pueden utilizar para calcular

sus propios estadísticos de desempeño, según se describe en el apartado 7, o incluirse en los gráficos que se describen en el apartado 8, pero el informe mencionará cuantas mediciones reportaron y que resultados son menores que el número exigido por el programa.

## **6. Determinación de la desviación típica para la evaluación de la aptitud**

*Hay un conjunto de métodos, pero solo vamos a describir los que consideramos más factible utilizar. El coordinador tendrá la responsabilidad de la selección del mismo.*

### **6.1 A partir de los resultados de un experimento de precisión**

Si el método de medición que se va a utilizar en un programa de ensayos de aptitud está normalizado y se dispone de información, sobre la repetibilidad y reproducibilidad del método, la desviación típica para la evaluación de la aptitud  $\hat{\sigma}$  se puede calcular a partir de dicha información del siguiente modo:

Datos:

$\sigma_R$  desviación típica de la reproducibilidad

$\sigma_r$  desviación típica de la repetibilidad

Calcule la desviación típica entre laboratorio como:

$$\sigma_L = \sqrt{\sigma_R^2 - \sigma_r^2}$$

y luego calcule la desviación típica para la evaluación de la aptitud como:

$$\hat{\sigma} = \sqrt{\sigma_L^2 + (\sigma_r^2/n)}$$

donde n es el número de mediciones repetidas (réplicas), que deberá realizar cada laboratorio en una ronda del programa.

### **6.2 A partir de los datos obtenidos en una ronda de un ensayo de aptitud**

En este enfoque, la desviación típica  $\hat{\sigma}$  para la evaluación de la aptitud de los participantes de una ronda de programa se obtiene a partir de los resultados reportados por los participantes de la misma ronda. El número de laboratorios participantes para emplear en este enfoque deberá ser  $\geq 15$ . La desviación típica puede ser la desviación típica robusta de los resultados reportados por todos los participantes, calculada por el Algoritmo A del Anexo C (véase 5.6.1). También la desviación típica  $\hat{\sigma}$  se podrá calcular usando el método de eliminar los valores atípicos (véase 5.6.2).

En este contexto, el “resultado” reportado por un participante será el promedio de las n mediciones repetidas (réplicas) obtenidas por el participante en la ronda. Una desventaja de este enfoque es que el valor  $\hat{\sigma}$  puede variar sustancialmente de una ronda a otra, la cual dificulta el uso de valores Z por parte del laboratorio al definir tendencias que persistan en varias rondas.

Esta desventaja se puede eliminar en un programa establecido con el uso de un valor robusto acordado de las desviaciones típicas obtenidas en una serie de rondas, calculado con el Algoritmo S del Anexo C. También es factible utilizar la dócima de Cochran (Anexo D) para eliminar los valores atípicos y calcular la desviación típica promedio a partir de las desviaciones típicas restantes de las rondas, por la expresión:

$$\bar{s} = \sqrt{\frac{\sum s_i^2}{k}}$$

## 7. Cálculo de los estadísticos de desempeño.

Los resultados de ensayo de aptitud a menudo necesitan ser transformados en un estadístico de desempeño, para ayudar a la interpretación y para permitir la comparación con metas definidas. El objetivo es medir la desviación del resultado de ensayo de cada participante con el valor asignado que permita la comparación con criterios de desempeño.

Son indicados y explicados más abajo en orden creciente de transformación de los resultados de los participantes, los estadísticos de desempeño así como su interpretación.

### 7.1 Estimados del sesgo del laboratorio

#### 7.1.1 Cálculo e interpretación

Tenemos que  $x$  representa el resultado (o el promedio de los resultados), reportados por un participante en la medición de una característica del material de ensayo en una ronda de un programa de ensayo de aptitud.

Entonces un estimado del sesgo  $D$  del laboratorio cuando se mide dicha característica, se calcula como:

$$D = x - X$$

donde  $X$  es el valor asignado.

La diferencia simple entre el resultado del participante y el valor asignado puede ser adecuado para determinar el desempeño y es comprendida más fácilmente por los participantes.

#### Interpretación de los sesgos de los laboratorios.

Los sesgos de los laboratorios se interpretarán mediante lineamientos equivalentes a los establecidos para los valores  $Z$ , o sea, cuando un participante reporta un sesgo del laboratorio superior a  $3,0 \hat{\sigma}$  o inferior a  $-3,0 \hat{\sigma}$ , entonces se considerará que el resultado ofrece una “señal de acción” es decir el sesgo es significativo y el resultado del ensayo se considera que no pertenece a la población de los resultados de ensayo.

De igual modo, se considerará que un sesgo del laboratorio superior a  $2,0 \hat{\sigma}$  o inferior a  $-2,0 \hat{\sigma}$  ofrece una “señal de advertencia”. Una única “señal de acción” en una ronda, o dos “señales de advertencias” en rondas sucesivas, se tomarán como evidencia que ha ocurrido una anomalía que requiere investigación.

#### Ejemplo:

La tabla 5 muestra los resultados de la aplicación de este método a los datos de la tabla 2

Tabla 5. Estimados del sesgo de laboratorio de las concentraciones de un anticuerpo I<sub>g</sub> E.

Laboratorio	Diferencia $x - X$ (sesgo)
A	0,27
B	-2,74
C	0,87
D	4,57
E	2,37
F	1,47
G	-0,63
H	-1,65
I	3,17
J	1,07
K	-2,93
L	-0,23
M	2,77
N	-4,03
O	-0,18
P	-8,85 W
Q	-2,64
R	-4,08
S	0,77
T	-0,13
U	5,27
V	-1,32
W	-0,53
X	2,57
Y	-0,93
Z	5,04
a	-2,56
3,0 $\hat{\sigma}$	9,12
2,0 $\hat{\sigma}$	6,08
-2,0 $\hat{\sigma}$	-6,08
-3,0 $\hat{\sigma}$	-9,12

Nota1 W = señal de advertencia

Nota 2. Los sesgos de laboratorios se obtuvieron de los datos de la tabla 2, usando el promedio robusto como el valor asignado. Los límites de acción y advertencia se calcularon a partir de la desviación típica robusta usándola como la desviación típica para la evaluación de la aptitud.

### 7.1.2 Evaluación del desempeño

De cada participante

Diferencia  $|x - X| \leq 2 \hat{\sigma} =$  satisfactorio

$2 \hat{\sigma} <$  diferencia  $|x - X| < 3 \hat{\sigma} =$  cuestionable

Diferencia  $|x - X| > 3 \hat{\sigma} =$  no satisfactorio

## 7.2 Diferencias porcentual.

### 7.2.1 Calculo e interpretación.

Con la notación de 7.1, las diferencias porcentuales de cada participante se calculan como sigue:

$$D_{\%} = 100 (x - X) / X$$

#### **Interpretación de las diferencias porcentuales.**

Las diferencias porcentuales se interpretarán mediante lineamientos, equivalentes a los establecidos para los sesgos de laboratorio del 7.1 o sea cuando un participante reporta un resultado que produzca una diferencia porcentual superior a  $300 \hat{\sigma} / X$  o inferior a  $-300 \hat{\sigma} / X$ , entonces se considera que el resultado ofrece una señal de acción. De igual modo, se considerará que una diferencia porcentual superior a  $200 \hat{\sigma} / X$  o inferior a  $-200 \hat{\sigma} / X$  ofrece una "señal de advertencia". Una única señal de acción o "señales de advertencia" en dos rondas sucesivas se tomarán como evidencia de que ha ocurrido una anomalía que requiere investigación.

#### **Ejemplo:**

La tabla 6 muestra los resultados de la aplicación de este método a los datos de la tabla 2.

Este método tiene la ventaja de trabajar con el porcentaje del sesgo  $(x - X)$  en relación al valor asignado y relacionarlo con la tolerancia del requisito técnico.

Tabla 6. Diferencias porcentual de laboratorio de la concentración de un anticuerpo I<sub>g</sub> E

Laboratorio	Diferencial porcentual %
A	2
B	-25
C	8
D	41
E	21
F	13
G	-6
H	-15
I	29
J	10
K	-27
L	-2
M	25
N	-37
O	-2
P	-80 W
Q	-24
R	-37
S	7
T	-1
U	48
V	-12
W	-5
X	23
Y	-8
Z	46
a	-23
300 $\hat{\sigma}/X$	82,7
200 $\hat{\sigma}/X$	55,1
-200 $\hat{\sigma}/X$	-55,1
-300 $\hat{\sigma}/X$	82,7
<p>Nota 1 W = Señal de advertencia  Nota 2 Las diferencias porcentual se obtuvieron de los datos de la tabla 2, usando el promedio robusto como el valor asignado. Los límites de acción y advertencia se calcularon a partir de la desviación típica robusta usándola como la desviación típica para la evaluación de la aptitud.</p>	

### 7.2.2 Evaluación del desempeño

De cada participante

$$|D_{\%}| \leq 200 \hat{\sigma} / X = \text{satisfactorio}$$

$$200 \hat{\sigma} / X < |D_{\%}| < 300 \hat{\sigma} / X = \text{cuestionable}$$

$$|D_{\%}| > 300 \hat{\sigma} / X = \text{no satisfactorio}$$

### 7.3 Clasificaciones y clasificaciones del porcentaje.

Con los resultados de  $p$  laboratorios en una ronda de un ensayo de aptitud, las clasificaciones son obtenidas asignando las clasificaciones de 1 al laboratorio que reporta el resultado más bajo, 2 al laboratorio que reporte el siguiente resultado más bajo y así sucesivamente hasta que el laboratorio que reporta el resultado más alto es asignado la clasificación de  $p$ . Si el ensayo se hace en varios mesurandos, las clasificaciones son asignadas en cada mesurando separadamente.

Si las clasificaciones son denotadas por  $i = 1, 2, \dots, p$  entonces las clasificaciones del porcentaje se calculan como  $100 \times (i - \frac{1}{2}) / p$ . En la tabla 7 se muestra un ejemplo de cómo se calculan las clasificaciones y las clasificaciones del porcentaje.

#### Interpretación de las clasificaciones y las clasificaciones del porcentaje.

La interpretación de las clasificaciones y las clasificaciones del porcentaje no implica la suposición de que los datos siguen una distribución de probabilidad particular, y su obtención no utiliza el valor asignado ni la desviación típica para la evaluación de la aptitud

De ahí que las clasificaciones y las clasificaciones del porcentaje constituyen un método sencillo para identificar a los laboratorios que reportan los resultados más extremos. Son útiles sobre todo en las primeras rondas de un programa de aptitud, cuando se puede utilizar para identificar a los laboratorios donde es más probable lograr mejoras del desempeño. El informe de desempeño mediante la clasificación de laboratorios en una tabla de acuerdo con su desempeño no es recomendable, en los ensayos de aptitud. Por tanto, la clasificación se debe utilizar solamente con extremo cuidado, ya que puede dar lugar a confusiones e interpretaciones erróneas.

Tabla 7. Clasificaciones y clasificaciones del porcentaje de un anticuerpo I<sub>g</sub> E

Laboratorio	Clasificación	Clasificación del porcentaje
A	16	57
B	5	17
C	18	65
D	25	91
E	21	76
F	20	72
G	11	39
H	8	28
I	24	87
J	19	69
K	4	13
L	13	46
M	23	83
N	3	9
O	14	50
P	1	2
Q	6	20
R	2	6
S	17	61
T	15	54
U	27	98
V	9	31
W	12	43
X	22	80
Y	10	35
Z	26	94
a	7	24

Nota: Los laboratorios con resultado más extremos son: P(2), R(6); N (9); U (98); Z (94) y D (91)

## 7.4 Valores Z

### 7.4.1 Cálculo e interpretación

Con la notación de 7.2, los valores Z se calculan como sigue:

$$Z = (x - X) / \hat{\sigma}$$

Donde  $\hat{\sigma}$  es la desviación típica para la evaluación de la aptitud. Este modelo puede ser usado cuando X y  $\hat{\sigma}$  se derivan de los resultados de los participantes, así como cuando X y  $\hat{\sigma}$  no se derivan de los resultados de (todos) los participantes.

#### Interpretaciones de los valores Z.

Cuando un participante reporta un resultado que produzca un valor superior a 3,0 o inferior a -3,0, entonces se considerará que el resultado ofrece una "señal de acción". De igual modo, se considerará que un valor Z superior a 2,0 ó inferior a -2,0 ofrece una

“señal de advertencia”. Una única “señal de acción” o “señales de advertencia” en dos rondas sucesivas se tomarán como evidencia de que ha ocurrido una anomalía que requiere investigación. Cuando se trata de unos pocos laboratorios (digamos menos de 10) puede que no se produzca ninguna señal.

En este caso, los métodos gráficos que combinan los valores de varias rondas darán indicaciones más útiles sobre el desempeño de los laboratorios que los resultados de rondas individuales.

**Ejemplo:**

La tabla 8 muestra los valores de Z mediante el promedio robusto y la desviación típica robusta que aparecen en la tabla 2.

Consideramos que este método es el más adecuado por razones obvias cuando el número de laboratorio participante es grande, siendo más fácil su interpretación y llegar a conclusiones estadísticas.

Recomendamos su utilización con preferencia a los demás métodos cuando sea posible su empleo.

Este método lo consideramos el más apropiado para los ensayos de aptitud sobre las características de los productos químicos donde se emplean métodos analíticos.

Tabla 8. Valores Z de un anticuerpo I<sub>g</sub> E

Laboratorio	Valor Z
A	0,09
B	-0,90
C	0,29
D	1,50
E	0,78
F	0,48
G	-0,21
H	-0,54
I	1,04
J	0,35
K	-0,96
L	-0,08
M	0,91
N	-1,33
O	-0,06
P	-2,91 W
Q	-0,87
R	-1,34
S	0,25
T	-0,04
U	1,73
V	-0,43
W	-0,17
X	0,85
Y	-0,31
Z	1,66
a	-0,84

Nota 1 W= señal de advertencia  
 Nota 2 Los valores Z en esta tabla se obtuvieron de los datos de la tabla 2, utilizando el promedio robusto como el valor asignado y utilizando la desviación típica robusta como las desviaciones típicas para la evaluación de la aptitud. Por tanto, la fórmula para el valor Z en este ejemplo es:

$$Z = (x - X) / \hat{\sigma} = (x - x^*) / s^*$$

#### 7.4.2 Evaluación del desempeño

De cada participante

$|Z| \leq 2$  = satisfactorio

$2 < |Z| < 3$  = cuestionable

$|Z| \geq 3 =$  no satisfactorio

## 7.5 Números $E_n$

*Normalmente usados en programas de comparaciones de mediciones (véase 4.2 de la NC-ISO/IEC 43-1) donde el valor asignado es provisto por un laboratorio de referencia, que podría ser la máxima autoridad de un país para la medición en cuestión. Los resultados de medición individuales se compararán con el valor asignado y la incertidumbre del laboratorio de referencia. El coordinador debe tomar en cuenta la incertidumbre de medición declarada de cada laboratorio que participa.*

El valor asignado y la incertidumbre típica del laboratorio de referencia se calcula utilizando el procedimiento 5.6.1 ó el 5.6.2 empleando 10 ó más resultados de ensayo (de las mediciones en paralelo). Consideramos que el procedimiento 5.6.2 (eliminando los valores atípicos) es el más apropiado por ser menos complejo.

La incertidumbre típica de los participantes no debe ser estimada a partir de la unidad de ensayo, por la afectación que podría sufrir la unidad de ensayo. La incertidumbre se calcula multiplicando por 3 la incertidumbre típica.

También es posible su utilización en el caso de un material de ensayo para la evaluación de desempeño de uno o más laboratorios a partir de un laboratorio de referencia, se utiliza cuando el número de laboratorio participante son pocos (<10)

### 7.5.1 Cálculo del estadístico de desempeño y su interpretación

El estadístico de desempeño se calcula como:

$$E_n = \frac{x - X}{\sqrt{U_{lab}^2 + U_{ref}^2}}$$

donde:

$X$  valor asignado del laboratorio de referencia

$x$  el resultado de ensayo

$U_{ref}$  incertidumbre ( $3 u_x$ ) del valor asignado donde  $u_x = \frac{\hat{\sigma}_{ref}}{\sqrt{p}}$

$U_{lab}$  incertidumbre ( $3 u_{lab}$ ) del resultado  $x$  de un participante

A diferencia de los valores críticos de 2,0 y 3,0 utilizados con los valores Z, se utiliza el valor crítico de 1,0 con  $E_n$  número.

Cuando  $|E_n| > 1$  indica señal de acción.

Eso sucede porque  $E_n$  números son calculados usando la incertidumbre en el denominador en vez de desviaciones típicas.

### 7.5.3 Evaluación del desempeño

De cada participante

$$|E_n| \leq 1 \text{ satisfactorio}$$

$$|E_n| > 1 \text{ no satisfactorio}$$

### 7.6 Valores $Z'$

Con la anotación del 7.4, el valor  $Z'$  se calcula como:

$$Z' = (x - X) / \sqrt{\hat{\sigma}^2 + u_X^2}$$

Donde  $u_X$  es la incertidumbre típica del valor asignado  $X$

#### Interpretación de los valores $Z'$ .

Los valores  $Z'$  se interpretarán igual que los valores  $Z$  (véase 7.4) y utilizando los mismos valores críticos de 2,0 y 3,0.

Nota: Los valores críticos que aparecen en 7.1 y 7.2 para la interpretación de los sesgos de laboratorio y de las diferencias del porcentaje se puede modificar de igual

modo sustituyendo  $q_c$  por  $\sqrt{\hat{\sigma}^2 + u_X^2}$

#### Uso de los valores $Z'$ .

Si se cumple el lineamiento de 4.2 que  $u_X \leq 0,3 \hat{\sigma}$  entonces se puede despreciar la incertidumbre del valor asignado, en caso contrario hay que tener en cuenta lo expresado en este apartado.

Si no se cumple el lineamiento de 4.2, la diferencia en magnitud entre los valores  $Z'$  y los valores  $Z$  puede ser tal que algunos valores  $Z$  excederán los valores críticos de 2,0 y 3,0 y por tanto emitir "señales de advertencia" o "señales de acción", mientras que los correspondientes valores  $Z'$  no excederán a estos valores críticos y por tanto no emitirán señales.

## 8. Método gráfico para combinar los valores de desempeño tras varias rondas de un programa de aptitud

### 8.1 Aplicaciones

Cuando se van a cambiar los valores  $Z$  tras varias rondas, el coordinador tendrá en cuenta la elaboración de gráficos, según se describe en 8.2. El uso de este gráfico, en los que se combinan los valores  $Z$  para varias rondas, puede permitir la identificación de tendencias u otras características de los resultados que no son evidentes cuando los valores  $Z$  para cada ronda se examinan por separado.

Nota: No se recomiendan el uso de “valores de corrida”, donde los valores obtenidos por un laboratorio se combinan tras varias rondas, pero sin mostrarse gráficamente. Un laboratorio podría tener un fallo que aparecería con el material de ensayo utilizado en una ronda, pero no en otras, y un valor de corridas podría ocultar este fallo.

## 8.2 Gráfico de control de Shewhart

Para elaborar este gráfico, los valores Z para un laboratorio se plotean como puntos individuales, con los límites de acción y de advertencia fijados en  $\pm 2,0$  y  $\pm 3,0$  según se ilustra en la tabla 9 y la figura.

El gráfico de control de Shewhart es un método eficaz para identificar problemas que causan grandes errores en los valores de Z. Las reglas para interpretar el gráfico de control de Shewhart cuando aparece una señal fuera de control son:

- un único punto cae fuera de los límites de acción ( $\pm 3,0$ );
- dos puntos sucesivos caen entre el límite de advertencia y el límite de acción de un mismo lado de la línea central (0,0);
- siete puntos sucesivos quedan en el mismo lado de la línea central (0,0).

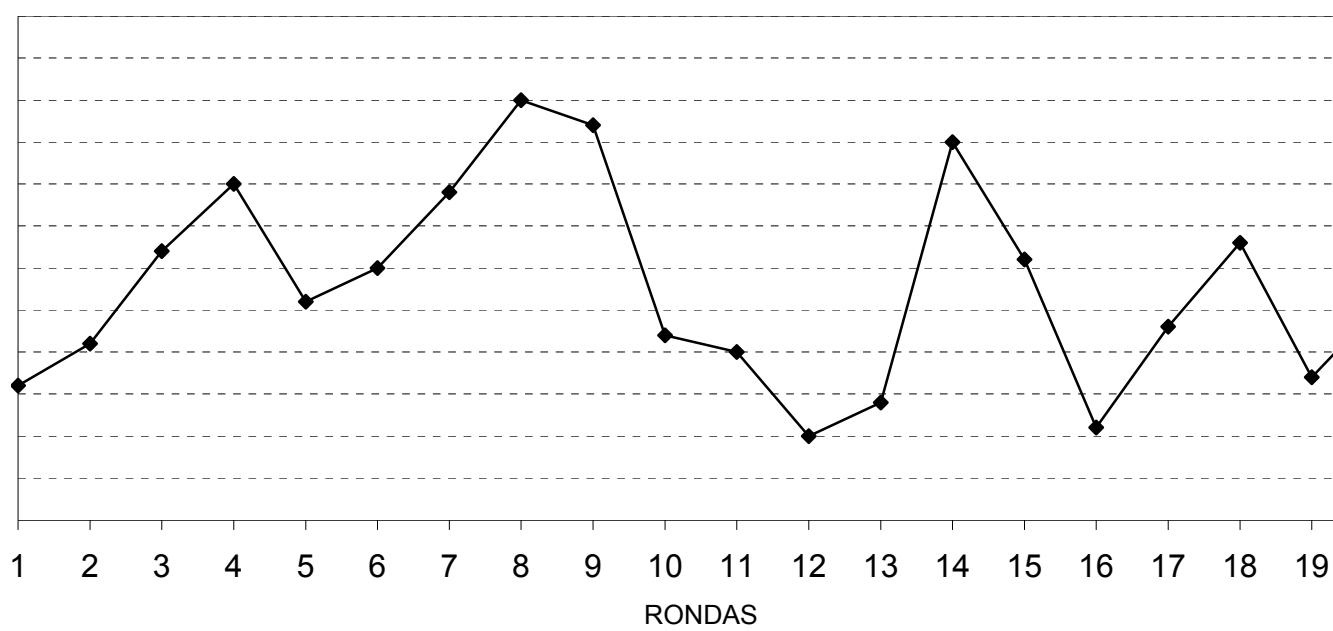
Cuando dicho gráfico de control de Shewhart refleja una señal de fuera de control, se iniciarán las medidas establecidas en 4.1.

### Ejemplo:

Concentración de un alergeno. Los datos se muestran en la tabla 9 y aparecen ploteados en la figura.

Tabla 9. Datos de 20 rondas de un ensayo de aptitud con determinación de la concentración de un alergeno.

Fecha de la ronda		Valor Z del alergeno
1991	Septiembre	-1,4
1991	Diciembre	-0,9
1992	Marzo	0,2
1992	Junio	1,0
1992	Septiembre	-0,4
1992	Diciembre	0,0
1993	Marzo	0,9
1993	Junio	2,0
1993	Septiembre	1,7
1993	Diciembre	-0,8
1994	Marzo	-1,0
1994	Junio	-2,0
1994	Septiembre	-1,6
1994	Diciembre	1,5
1995	Marzo	-0,1
1995	Junio	-1,9
1995	Septiembre	-0,7
1995	Diciembre	0,3
1996	Marzo	-1,3
1996	Junio	-0,4



**Fig. Gráfico de control de Shewart de los datos de un laboratorio de 20 rondas de un ensayo de aptitud.**

## 9. Programa de ensayo de muestras divididas

Un tipo de ensayo de aptitud el cual es usado a menudo por los clientes de los laboratorios incluidos algunos organismos regulatorios es la técnica de ensayo de muestras divididas. Las muestras divididas se utilizan cuando es necesario realizar una comparación detallada de dos laboratorios. Pudiendo utilizarse en los siguientes casos:

- si un laboratorio es operado por un proveedor y otra por un cliente y las dos organizaciones desean garantizar que los dos laboratorios estén en concordancia;
- comparación entre laboratorios que están siendo evaluados como proveedores potenciales o proveedores de servicios de ensayo;
- comparación entre un laboratorio que se desea evaluar con otro laboratorio de referencia.

Tales programas frecuentemente requieren retener suficiente material para resolver cualquier diferencia percibida entre los dos laboratorios involucrados, por medio de un análisis adicional por un laboratorio de tercera parte. El programa de ensayo de muestra dividida difiere al programa de ensayos interlaboratorio puesto que hay un número muy limitado de laboratorios participantes (a menudo dos)

Se obtienen muestras de varios materiales (procedentes de varios lotes) y cada muestra se divide en dos partes, y luego cada laboratorio obtiene un número (al menos dos) de determinaciones repetidas (réplicas) en cada submuestra. Luego se determinará si hay diferencia significativa entre las submuestras de las dos poblaciones (entre laboratorios) tanto en las medidas y las desviaciones típicas. Si hay diferencia significativa entre las desviaciones típicas es indicativo que hay diferencia entre la variabilidad (precisión) de ambos laboratorios.

### Procedimiento

- Se toman 10 ó más muestras preferentemente 15.
- Se confecciona la tabla de datos experimentales con los resultados pareados que proceden de los resultados de las dos poblaciones. Los resultados son obtenidos por n mediciones repetidas (réplicas) en cada submuestra.

La tabla 10 es un modelo como deben procesarse los datos para aplicar las dójimas de “comparación de dos medias para el caso de observaciones pareadas” y la dójima de “comparación de dos varianzas ó dos desviaciones típicas”

Tabla 10. Modelo de registro y procesamiento de los datos

Muestra	$x_i$	$y_i$	$d_i$	$d_i^2$
1	$x_1$	$y_1$	$d_1$	$d_1^2$
2	$x_2$	$y_2$	$d_2$	$d_2^2$
3	$x_3$	$y_3$	$d_3$	$d_3^2$
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
$k$	$x_k$	$y_k$	$d_k$	$d_k^2$
	$\sum x_i$	$\sum y_i$	$\sum d_i$	$\sum d_i^2$

donde:

$$n \geq 10$$

$x_i$  : resultado i-ésimo de las muestras del laboratorio A (promedio de dos réplicas)

$y_i$  : resultado i-ésimo de las muestras del laboratorio A (promedio de dos réplicas)

$$d_i = x_i - y_i$$

$k$  : número de muestras

- Se aplican:

### **Décima de comparación de dos medias para el caso de observaciones pareadas.**

Se aplica para conocer si hay diferencia significativa entre las medias de los dos laboratorios.

- Se procede a determinar si hay diferencia significativa entre las diferencias ( $d_i$ ) de los pares de resultados empleando la décima de Cochran, como se indica a continuación:
  - a) Los valores que son manifiestamente atípicos, ya sea por razones experimentadas, técnicas o evidentes, de la serie de las diferencias observadas ( $d_i$ ) deben ser eliminados. Continuar el proceso con los datos restantes.
  - b) Para detectar los resultados atípicos de las diferencias.

Siendo:  $k$  el número de pares y

$$g_i = d_i^2 / \sum d_i^2$$

**La décima conduce a eliminar aquellos pares de valores ( $x_i, y_i$ ) cuando los valores de  $g_i$  son superiores a los valores críticos  $g_{0,95}(\alpha = 5\%)$  ó  $g_{0,99}(\alpha = 1\%)$  según el nivel de confianza establecido, los cuales se dan en la tabla 11.**

Tabla 11. Valores críticos de la Cochran para  $\alpha = 5\%$  y  $\alpha = 1\%$

$k$	$g_{0,95}(\alpha = 5\%)$	$g_{0,99}(\alpha = 1\%)$
10	0,60	0,72
12	0,54	0,65
15	0,47	0,57

**Cálculos:**

$$\bar{d} = \frac{1}{k} (\sum x_i - \sum y_i) = \frac{1}{k} \sum d_i =$$

$$s_d^2 = \frac{1}{k-1} \left[ \sum d_i^2 - \frac{1}{k} (\sum d_i)^2 \right] =$$

$$s_d = \sqrt{s_d^2} =$$

$$A_2 = \left[ \frac{t_{1-\alpha/2}(v)}{\sqrt{k}} \right] s_d =$$

Los valores de  $\frac{t_{1-\alpha/2}(v)}{\sqrt{k}}$  se dan en la tabla 12

## Resultado

### Décima bilateral:

La hipótesis de que la media de la población de las diferencias es igual a la hipótesis nula = 0 se rechaza si

$$|\bar{d}| > A_2$$

Esto es indicativo de que la media de la población X (laboratorio A) es significativamente diferente a la media de la población Y (laboratorio B)

Cuando  $\bar{d}$  es positiva y la hipótesis nula, se rechaza, la media de la población X es mayor a la de Y y cuando  $\bar{d}$  es negativa y la hipótesis nula se rechaza, la media de la población Y es mayor a la de X.

NOTA: En este trabajo se recomienda utilizar el nivel de significación  $\alpha = 5\%$

Tabla 12. Valores de la relación  $t_{1-\alpha/2}(v)/\sqrt{k}$  para  $v = k - 1$

v = k - 1	Décima bilateral	
	$\alpha = 0,05$	$\alpha = 0,01$
	$t_{0,975}/\sqrt{k}$	$t_{0,995}/\sqrt{k}$
8	0,769	1,118
9	0,715	1,028
10	0,672	0,956
11	0,635	0,897
12	0,604	0,847
13	0,577	0,805
14	0,554	0,769
15	0,533	0,737

Décima de comparación de dos varianzas ó de dos desviaciones típicas.

Se aplica para conocer si hay diferencia significativa entre las desviaciones típicas de los dos laboratorios

## Cálculos

Se calculan las varianzas de los dos laboratorios por las expresiones siguientes:

Laboratorio A

$$s_x^2 = \frac{\sum (x_i - \bar{x})^2}{k-1} =$$

Laboratorio B

$$s_y^2 = \frac{(\sum y_i - \bar{y})^2}{k-1} =$$

y a partir de estos se hallan los valores de:

$$\frac{s_x^2}{s_y^2} = F_{1-\alpha/2}(k_1, k_2) = \frac{1}{F_{1-\alpha/2}(k_2, k_1)} =$$

Los valores de F se obtienen a partir de la tabla 13. En este caso  $k_1 = k_2$

## Resultados

Comparación de las varianzas de dos poblaciones. La hipótesis de la igualdad de la varianza se rechaza si:

$$\frac{s_x^2}{s_y^2} < \frac{1}{F_{1-\alpha/2}(k_2, k_1)} \quad \text{o} \quad \frac{s_x^2}{s_y^2} > F_{1-\alpha/2}(k_1, k_2)$$

Nota: Si  $s_x^2/s_y^2$  es menor que el límite inferior se considera que la varianza (precisión) del laboratorio A es menor a la del Laboratorio B y lo contrario, si es mayor al límite superior.

Tabla 13. Valores superiores de F. Valores de  $F_{0,975}(k_1, k_2)$ ,  $\alpha = 0,05$

$k_1 \backslash k_2$	8	9	10	11	12	13	14	15
8	4,99	4,90	4,8	4,76	4,71	4,67	4,63	4,60
9	4,53	4,43	4,36	4,30	4,24	4,20	4,16	4,13
10	4,20	4,10	4,03	3,96	3,91	3,87	3,83	3,80
11	3,95	3,85	3,78	3,72	3,66	3,62	3,58	3,55
12	3,76	3,66	3,59	3,53	3,47	3,43	3,39	3,36
13	3,61	3,51	3,44	3,37	3,32	3,28	3,21	3,18
14	3,48	3,39	3,31	3,25	3,20	3,15	3,12	3,08
15	3,38	3,29	3,21	3,15	3,09	3,05	3,01	2,98

Ejemplo: Para  $\alpha = 5\%$

$$F_{0,975}(8,10) = 4,20 \quad F_{0,025}(8,10) = \frac{1}{F_{0,975}(10,8)} = \frac{1}{4,82} = 0,207$$

## 10. Programas de ensayo con valores conocidos

Estos programas se caracterizan porque se conoce el valor asignado empleando uno de los siguientes métodos: Formulación (véase 5.2), valores de referencia certificados (véase 5.3) y valores de referencia (véase 5.4) y también se conoce la desviación para la evaluación de la aptitud ( $\hat{\sigma}$ ) mediante el enfoque. “A partir de los resultados de un experimento de precisión”, es decir, cuando se conoce la repetibilidad y la reproducibilidad del método de ensayo (véase 6.1) o cuando se prescribe o se fija  $\hat{\sigma}$ . Tales programas de ensayos de aptitud no necesitan la participación de múltiples laboratorios ( $\geq 10$ ) y se emplea el estadístico de desempeño “Valores Z” (véase 7.4)

## 11. Programa de comparación de mediciones

*En este programa de aptitud (véase 4.2 de la NC-ISO/IEC Guía 43-1) se emplea el estadístico de desempeño “Números  $E_n$ ” y se tendrá en cuenta lo establecido en 7.5.*

### 11.1 Aplicación de métodos estadísticos en el programa de comparación de mediciones.

- a) *Se determina el valor asignado y la incertidumbre típica del laboratorio de referencia a partir de 10 ó más resultados de ensayo ( de dos mediciones simples en paralelo) sobre la unidad de ensayo de acuerdo con el procedimiento 5.6.1 ó 5.6.2 sustituyendo donde dice resultado de un participante por el resultado de dos mediciones en paralelo, consideramos que el procedimiento 5.6.2 es el más apropiado por ser más sencillo.*
- b) Los laboratorios participantes deberán informar la incertidumbre típica para:  $n = 2$ , la cual se puede estimar a partir de 10 resultado de ensayos de dos mediciones simples (réplicas), en paralelo, en condiciones de repetibilidad correspondiente a una unidad de ensayo de un mismo nivel de expresión de la característica objeto de estudio. Se determina la desviación típica de los resultados de un participante de acuerdo con el procedimiento 5.6.2.

Esta desviación típica ( $s = \hat{\sigma}$ ) corresponde a la incertidumbre típica de un resultado del laboratorio participante ( $u_{lab}$ ) y si se multiplica por 3 se tiene la incertidumbre del resultado  $X$  ( $U_{lab}$ ). También si se conoce la  $s_r$  y  $s_R$  del laboratorio participante se puede calcular la incertidumbre típica ( $u_{lab}$ ) de un resultado por la expresión:

$$u_{lab} = \sqrt{s_L^2 + \frac{s_r^2}{n}}$$

donde:

n número de mediciones simples (réplicas)

$$s_L^2 = s_R^2 - s_r^2 / n$$

- c) Se aplica el estadístico de desempeño “Número  $E_n$ ” (véase 7.5)
- d) Se compara el estadístico de desempeño  $E_n$  de varias rondas para conocer el mejoramiento del desempeño de cada laboratorio.

**Nota:** También es factible utilizar un procedimiento similar para la evaluación de desempeño de uno o más laboratorios mediante muestras idénticas de un material de ensayo cuando el número de laboratorios participantes son pocos ( $\leq 10$ ). En este caso el resultado es sobre una muestra idéntica donde se efectúan dos réplicas.

## **12. Pasos a seguir para la aplicación de los métodos estadísticos en los programas de ensayos interlaboratorio.**

1. Definir la identidad de los laboratorios participantes y su número
2. Selección y aplicación del método para determinar el valor asignado y su incertidumbre típica (véase 5)
3. Selección y aplicación del método para la determinación de la desviación típica para la evaluación de la aptitud (véase 6)
4. Determinar la homogeneidad y la estabilidad de las muestras (véase 4.4 y Anexo B)
5. Determinar si la incertidumbre típica del valor asignado satisface el requisito que nos permite despreciarla en la interpretación de los resultados del ensayo de aptitud, en el caso contrario para hacer los ajustes pertinentes en los cálculos de los estadísticos de desempeño (véase 4.2 y 7.6)
6. Determinar el número de réplicas (mediciones repetidas) véase 4.3
7. Selección y aplicación del estadístico de desempeño de acuerdo con las condiciones existentes (véase 7)
8. Aplicación de métodos gráficos para combinación de los valores de desempeño tras varias rondas (véase 8)

Ejemplo:

**En un ensayo de aptitud de la característica de un producto químico empleando un método de ensayo normalizado se tomarán las siguientes decisiones:**

1. Utilizar el tipo “programas de ensayo interlaboratorio” de acuerdo con la naturaleza de la unidad de ensayo, el método en uso, y el número de laboratorios participantes con la finalidad de evaluar el desempeño de los laboratorios participantes (véase 4.1 de la NC-ISO/IEC Guía 43-1)
2. Participación de 15 laboratorios, los cuales reúnen las características apropiadas, a las cuales se definieron su identidad y se le dio un código de identificación para mantener la confidencialidad
3. Seleccionar y aplicar el método de “Valores de consenso de los laboratorios participantes” (véase 5.6) para determinar el valor asignado y su incertidumbre típica por ser el que más se ajusta a las condiciones existentes y entre las dos variantes, el enfoque “Eliminando los valores atípicos” (véase 5.6.2)
4. Seleccionar y aplicar el método “A partir de los datos obtenidos en una ronda de un ensayo de aptitud” para determinar la desviación típica para la evaluación de la aptitud (véase 6.2)
5. Como el producto químico es un líquido y se conoce históricamente su homogeneidad y su estabilidad en el tiempo de duración del ensayo de aptitud, no es necesario realizar los ensayos pertinentes (véase 4.4 y Anexo B).
6. Cuando se emplea este programa con las alternativas expresadas anteriormente al aplicar el requisito de la incertidumbre típica del valor asignado:

$$u_x \leq 0,3 \hat{\sigma}$$

cuando se toma la decisión 4 se tiene que la desviación típica para la evaluación de la aptitud corresponde a la desviación típica de la población de los resultados de los laboratorios participantes estimada por la fórmula clásica, en este caso se tiene:

$$u_x = \frac{s}{\sqrt{p}}$$

De acuerdo a la decisión 2 se tiene que  $p = 15$ , haciendo las sustituciones pertinentes en la fórmula se tiene:

$$u_x = \frac{s}{\sqrt{15}} = \frac{s}{3,873} = 0,26 s < 0,3 s$$

Se cumple el requisito, por lo tanto se permite despreciar la incertidumbre típica en el estadístico de desempeño (véase 7.6 “Valores Z’ ”)

7. En la primera ronda se inicia con dos réplicas, si en ellas no se cumple el requisito:

$$\sigma_r / \sqrt{n} \leq 0,3 \hat{\sigma}$$

Se determina  $n$  para la próxima ronda (véase 4.3)

Si no se conoce  $\sigma_r$ , se utiliza  $s_r$  estimada según 5.6.2

8. Se selecciona y aplica el estadístico de desempeño “Valores Z” (véase 7.4)

9. Cuando hay una “señal de acción” y por lo tanto la evaluación de desempeño es “no satisfactorio” se le informa al laboratorio pertinente para que tome las medidas correctivas pertinentes. En las próximas rondas se observa si el laboratorio ha mejorado su desempeño

10. Cuando se tienen los valores de desempeño (valores Z) de 10 o más rondas se aplica el “Gráfico de Control de Shewhart (véase 8.2)

## ANEXO A

### Símbolos

$g$	número de muestras ensayadas en una comprobación de la homogeneidad
$n$	número de mediciones repetidas (réplicas) por muestra
$p$	número de laboratorios que toman parte en una ronda de un ensayo de aptitud
$s_L$	desviación típica entre las muestras
$s_{\bar{x}}$	<b>desviación típica de los promedios de las muestras</b>
$s^*$	<b>desviación típica robusta (desviación típica calculada con un algoritmo robusto)</b>
$s_r$	<b>estimación de las desviaciones típicas de la repetibilidad</b>
$s_R$	<b>estimación de la desviación típica de la reproducibilidad</b>
$u_x$	<b>incertidumbre típica del valor asignado</b>
$U$	<b>incertidumbre (3 <math>u</math>)</b>
$x$	<b>resultado de la medición</b>
$x^*$	<b>promedio robusto (promedio calculado con un algoritmo robusto)</b>
$Z$	<b>valor utilizado para la evaluación de la aptitud</b>
$X$	<b>valor asignado para la evaluación de la aptitud</b>
$\sigma_L$	<b>desviación típica entre laboratorios</b>
$\hat{\sigma}$	<b>desviación típica estimada para la evaluación de la aptitud</b>
$\sigma_r$	<b>desviación típica de la repetibilidad</b>
$\sigma_R$	<b>desviación típica de la reproducibilidad</b>
$\delta$	<b>factor utilizado para calcular <math>\hat{\sigma}</math> por análisis robusto</b>
$E_n$	<b>estadístico de desempeño usado en programas de comparaciones de mediciones.</b>

## ANEXO B

### Comprobación de la homogeneidad y la estabilidad entre las muestras

#### B.1 Procedimiento para comprobar la homogeneidad

- Escoja un laboratorio donde realizar la comprobación de la homogeneidad. Escoja un método de medición para utilizar en la comprobación y la característica que va a medir en el material de ensayo, que será sensible a la homogeneidad entre las muestras. Así, en un material granulado, escoja una característica totalmente liberada (por ejemplo, el porcentaje que pasa por un tamiz con determinado tamaño de apertura). Si la medición es una proporción, escoja una característica que esté presente como una proporción pequeña.
- Prepare y envase las muestras para una ronda del ensayo de aptitud, asegurando que haya muestras suficientes para los participantes y para la comprobación de la homogeneidad y la estabilidad.
- Seleccione un número  $g$  de muestras aleatorias, donde  $g \geq 10$
- Prepare dos porciones de ensayo de cada muestra utilizando técnicas apropiadas al material de ensayo con el fin de minimizar las diferencias entre dichas porciones de ensayo
- Tomando las 2  $g$  porciones de ensayo en orden aleatoria, obtenga un resultado de medición con cada una, completando toda la serie de medición en condiciones de repetibilidad. Elimine los resultados atípicos aplicando la dócima de Cochran y las dócimas de Grubbs
- Calcule el promedio general  $\bar{x}$ , la desviación típica de repetibilidad  $S_r$  y la desviación típica entre las muestras  $S_L$  como se muestra en B.3 de este Anexo.

#### B.2 Criterio de evaluación para una comprobación de la homogeneidad.

Compare la desviación típica entre las muestras  $S_L$  con la desviación típica  $\hat{\sigma}$  para la evaluación de la aptitud. Se pueden considerar las muestras adecuadamente homogéneas si:

$$s_L \leq 0,3 \hat{\sigma}$$

Si no se cumple este criterio, el coordinador tendrá en cuenta las siguientes posibilidades:

- Examinar el procedimiento de preparación de la muestra para ver si es posible mejorarlo**
- Distribuir un número de muestras a cada participante en el programa de aptitud y pedir que se obtenga un resultado de medición por cada muestra.**

La heterogeneidad de las muestras aumentará la desviación típica de la repetibilidad hasta

$$\sigma_{r1} = \sqrt{\sigma_r^2 + s_L^2}$$

Use  $\sigma_{r1}$  en vez de  $\sigma_r$  en el lineamiento para escoger el número de mediciones repetidas (réplicas). En la fórmula  $\sigma_r / \sqrt{n} \leq 0,3 \hat{\sigma}$  en 4.3

c) Incluya la desviación típica entre las muestras en la desviación típica para el ensayo de aptitud, calculando  $\hat{\sigma}$  como:

$$\hat{\sigma} = \sqrt{\hat{\sigma}_1^2 + s_L^2}$$

donde  $\hat{\sigma}_1$  es la desviación típica para el ensayo de aptitud que no permite heterogeneidad en las muestras.

### B. 3 Fórmulas para la comprobación de la homogeneidad .

Los datos para la comprobación de la homogeneidad están representados por:

$$x_{t k}$$

Donde:

$t$  representa a la muestra (  $t = 1,2, \dots,g$ )

$k$  representa a la porción de ensayo ( $k = 1,2$ ).

Definir los promedios de las muestras como:

$$\bar{x}_t = x_{t,1} + x_{t,2} / 2,0$$

y los recorridos entre las porciones de ensayo como:

$$R_t = |x_{t,1} - x_{t,2}|$$

Calcule el promedio general

$$\bar{x} = \sum \bar{x}_t / g$$

la desviación típica de los promedios de las muestras

$$s_{\bar{x}} = \sqrt{\sum (\bar{x}_t - \bar{x})^2 / (g - 1)}$$

y la desviación típica de la repetibilidad

$$s_r = \sqrt{\sum W_t^2 / (2g)}$$

donde las sumatorias están entre las muestras (  $t = 1, 2, \dots, g$  )

Finalmente, calcule la desviación típica entre las muestras

$$s_L = \sqrt{s_{\bar{x}}^2 - (s_r^2 / 2)}$$

La desviación típica de la reproducibilidad será:

$$s_R = \sqrt{s_r^2 + s_L^2}$$

Este procedimiento nos sirve también para estimar las desviaciones típicas de la repetibilidad y reproducibilidad interlaboratorios, aunque estos estimados están expuestos a grandes errores por ser el número de datos pequeño. Se puede lograr unos estimados adecuados aumentando el número de muestras a 15.

Los datos se registran y procesan de acuerdo a la tabla 14.

Tabla 14. Registro y procesamiento de los datos

Número de la muestra	Resultado de medición		Promedio	Recorrido
1	$x_{1,1}$	$x_{1,2}$	$\bar{x}_1$	$R_1$
2	$x_{2,1}$	$x_{2,2}$	$\bar{x}_2$	$R_2$

3	$x_{3,1}$	$x_{3,2}$	$\bar{x}_3$	$R_3$
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
t	$x_{t,1}$	$x_{t,2}$	$\bar{x}_t$	$R_t$
			$\bar{x}$	$R$

#### B.4 Procedimiento para la comprobación de la estabilidad

- Utilice para la comprobación de la estabilidad el mismo laboratorio para la comprobación de la homogeneidad. Utilice el mismo método de medición y mida la misma característica del material de ensayo.
- Permitir un margen de tiempo entre los ensayos de homogeneidad y los ensayos de estabilidad similar al tiempo que le toma ensayar las muestras a los participantes en el ensayo de aptitud.
- Seleccione un g de muestras aleatorias, donde  $g \geq 3$ .
- Prepare dos porciones de ensayo a partir de cada muestra, utilizando las mismas técnicas que en la comprobación de la homogeneidad.
- Tomando las 2 g porciones de ensayo en orden aleatoria, obtenga un resultado de medición  $y_{t,k}$  en cada una completando toda la serie de mediciones en condiciones de repetibilidad.
- Calcule el promedio general  $\bar{y}$  de las mediciones obtenidas en el ensayo de estabilidad.

#### B.5 Criterio de evaluación para la comprobación de la estabilidad

Compare el promedio general de las mediciones obtenidas en la comprobación de la homogeneidad con el promedio general de los resultados obtenidos en la comprobación de la estabilidad. Las muestras se pueden considerar como adecuadamente estables si:

$$\left| \bar{x} - \bar{y} \right| \leq 0,3 \hat{\sigma}$$

Si no se cumple este criterio, examine la preparación de la muestra y los procedimientos de almacenamiento para ver si es posible lograr mejoras.

#### B.6 Ejemplo: Cobre en la harina de soya ( $\mu_g / g$ )

En un programa de ensayos de aptitud en el cual se va a utilizar harina de soya como material de ensayo, la desviación típica para la evaluación de la aptitud se ha fijado en  $\hat{\sigma} = 1,1 \mu_g / g$  según el criterio  $s_L \leq 0,3 \hat{\sigma}$  que aparece en B.2, la desviación típica entre las muestras no debe ser mayor que:

$$0,3 \times 1,1 = 0,330 \mu_g / g$$

Para una comprobación de la homogeneidad, fueron seleccionados aleatoriamente 12 de las muestras de harina de soya preparadas para el programa de ensayos de aptitud, y se determinó el contenido de cobre de dos porciones de ensayo de cada muestra. Los datos aparecen en la tabla 15 junto con los promedios de las muestras y los recorridos entre las porciones de ensayo.

Se puede utilizar las fórmulas que aparecen B.3 para calcular los siguientes valores:

Promedio general  $\bar{x} = 10,02 \mu_g / g$

desviación típica de los promedios de las muestras  $s_{\bar{x}} = 0,340 \mu_g / g$

desviación típica de la repetibilidad  $s_r = 0,246 \mu_g / g$

y de aquí la desviación típica entre las muestras  $s_L = 0,282 \mu_g / g$

Este valor es menor que  $0,330 \mu_g / g$  de modo que podemos llegar a la conclusión de que las muestras son adecuadamente homogéneas para el uso en el programa de ensayo de aptitud.

Tabla 15. Resultados de medición para la comprobación de la homogeneidad, cobre en harina de soya ( $\mu_g / g$ )

Número de muestras $t$	Porción de ensayo 1	Porción de ensayo 2	Promedio de muestra $\bar{x}_t$	Recorrido entre las porciones de ensayo $R_t$
1	10,5	10,4	10,45	0,1
2	9,6	9,5	9,55	0,1
3	10,4	9,9	10,15	0,5
4	9,5	9,9	9,70	0,4
5	10,0	9,7	9,85	0,3
6	9,6	10,1	9,85	0,5
7	9,8	10,4	10,10	0,6
8	9,8	10,2	10,00	0,4
9	10,8	10,7	10,75	0,1
10	10,2	10,0	10,10	0,2
11	9,8	9,5	9,65	0,3
12	10,2	10,0	10,10	0,2

## ANEXO C

### Análisis robusto

#### C.1 Análisis robusto. Algoritmo A.

*Este algoritmo produce valores robustos del promedio y la desviación típica de los datos a los que se aplican donde disminuye los efectos de los valores extremos.*

El “promedio robusto” y la “desviación típica robusta” se deben interpretar en este trabajo como los estimados de la media de la población y de la desviación típica de la población calculados por medio de un algoritmo robusto.

Denotar los elementos  $p$  de los datos, clasificados en orden ascendente, como:

$$x_1, x_2, \dots, x_j, \dots, x_p$$

Denotar el promedio robusto y la desviación típica robusta de estos datos por  $x^*$  y  $s^*$

Calcular los valores iniciales para  $x^*$  y  $s^*$  como:

$$x^* = \text{mediana de } x_i \quad (i = 1, 2, \dots, p)$$

$$s^* = 1,483 \text{ mediana de } |x_i - x^*| \quad (i = 1, 2, \dots, p)$$

Actualizar los valores de  $x^*$  y  $s^*$  como sigue. Calcular:

$$\delta = 1,5 s^*$$

Por cada  $x_i$  ( $i = 1, 2, \dots, p$ ), calcular

$$x_i^* = \begin{cases} x^* - \delta & \text{si } x_i < x^* - \delta \\ x^* + \delta & \text{si } x_i > x^* + \delta \\ x_i & \text{de otra manera} \end{cases}$$

Calcular los nuevos valores de  $x^*$  y  $s^*$  a partir de:

$$x^* = \sum x_i^* / p$$

$$s^* = 1,134 \sqrt{\sum (x_i^* - x^*)^2 / (p - 1)}$$

Los estimados robustos  $x^*$  y  $s^*$  se puede obtener mediante cálculos repetitivos, es

decir actualizando los valores de  $x^*$  y  $s^*$  varias veces hasta que el proceso converja .

La convergencia se puede asumir cuando no haya cambios entre una repetición y otra en la tercera cifra significativa de la desviación típica robusta y de la cifra equivalente en el promedio robusto. Este es un método fácil de programar en una computadora.

Ejemplo: Véase 5.6

#### C.2 Algoritmo S

Este algoritmo se aplica a las desviaciones típicas (o recorridos). Produce un valor acumulado robusto de las desviaciones típicas o de los recorridos a los que se aplica.

Denotar los elementos  $p$  de los datos, clasificados, en orden ascendente, como:

$$w_1, w_2, \dots, w_i, \dots, w_p$$

(Estos pueden ser recorridos o desviaciones típicas). Denotar el valor acumulado

robusto como  $w^*$  y los grados de libertad asociados a cada  $w_i$  como  $v$  (Si  $w_i$  es

un recorrido,  $v = 1$ . Si  $w_i$  es la desviación de  $n$  resultado de ensayo,  $v = n - 1$  Obtener los valores de  $\xi$  y  $\eta$  requeridos por el algoritmo de la tabla 16.

Calcular un valor inicial para  $W^*$  como:

$$w^* = \text{mediana de } w_i (i=1,2, \dots, p)$$

Actualizar el valor de  $w^*$  como sigue. Calcular:

$$\psi = \eta \cdot w^*$$

Por cada  $w_i (i=1,2, \dots, p)$  calcular:

$$W_i^* = \begin{cases} \psi & \text{si } W_i > \psi \\ W_i & \text{de otra manera} \end{cases}$$

Calcular el nuevo valor de  $W^*$  a partir de:

$$w^* = \xi \sqrt{\sum (w_i^*)^2 / p}$$

Este estimado robusto  $W^*$  se puede obtener mediante cálculos repetitivos actualizando el valor de  $W^*$  varias veces hasta que el proceso converja. La convergencia se puede asumir cuando no haya cambios entre una repetición y otra en la tercera cifra significativa del  $W^*$ . Esto es un método fácil de programar en una computadora.

Tabla 16. Factores requeridos para el análisis robusto: Algoritmo S

Grado de libertad	Factor límite	Factor de ajuste
$v$	$\eta$	$\xi$
1	1,645	1,097
2	1,517	1,054
3	1,444	1,039
4	1,395	1,032
5	1,359	1,027
6	1,332	1,024
7	1,310	1,021
8	1,292	1,019
9	1,277	1,018
10	1,264	1,017

## ANEXO D

### Décimas de Grubbs (décima simple de Grubbs) y de Cochran

#### D.1.- Décima de Grubbs (Décima simple de Grubbs)

Detección de una única observación aberrante (atípica)

1. Se ordena en orden creciente el conjunto de datos  $x_i$  objeto de análisis, para determinar si la observación más alejada es incompatible con el resto

$$x_1 \leq x_2 \leq \dots \leq x_n$$

2. Se calcula la media muestral y la desviación típica muestral.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$$

3. Se calculan los estadísticos para la observación más grande y para la observación más pequeña.

$$G_n = \frac{(x_n - \bar{x})}{s} \quad G_1 = \frac{(\bar{x} - x_1)}{s}$$

Esta d cima se puede aplicar a las medias de muestras sustituyendo la observaci n ( $x$ ), la media muestral ( $\bar{x}$ ) y la desviaci n t pica muestral ( $s$ ) por la media de las observaciones ( $\bar{\bar{x}}$ ), la media de las medias ( $\bar{\bar{x}}$ ) y la desviaci n t pica de las medias ( $s_{\bar{x}}$ ), respectivamente, como se muestra a continuaci n:

$$\bar{x}_1 \leq \bar{x}_2 \leq \dots \leq \bar{x}_n$$

$$\bar{\bar{x}} = \frac{1}{n} \sum_{i=1}^n \bar{x}_i \quad s_{\bar{x}} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (\bar{x}_i - \bar{\bar{x}})^2}$$

$$G_n = \frac{(\bar{x}_n - \bar{\bar{x}})}{s_{\bar{x}}} \quad G_1 = \frac{(\bar{\bar{x}} - \bar{x}_1)}{s_{\bar{x}}}$$

4. Los resultados obtenidos de los estadísticos se comparan con sus valores cr ticos (valores l mite) para el tama o de muestra ( $n$ ) dado en la tabla 1.
5. Criterios para la aceptaci n o rechazo:
  - a) Si el valor estadístico ( $G$ ) es menor o igual al de su valor cr tico para un  $\alpha$  del 5% se acepta como correcto.
  - b) Si el valor estadístico es mayor al de su valor cr tico para un  $\alpha$  del 5% y menor o igual al de su valor cr tico para un  $\alpha$  del 1% se considera dudoso y se trata de localizar la causa asignable, si se determina se rechaza y en el caso contrario se acepta.
  - c) Si el valor del estadístico es mayor al de su valor cr tico para un  $\alpha$  del 1%, se considera estadisticamente incompatible y se rechaza.
6. Repita la d cima de Grubbs tantas veces como sea necesario

Tabla 17- Valores cr ticos para la d cima de Grubbs

n	$\alpha = 1\%$	$\alpha = 5\%$
3	1,155	1,155

4	1,496	1,481
5	1,764	1,715
6	1,973	1,887
7	2,139	2,020
8	2,274	2,126
9	2,387	2,215
10	2,482	2,290
11	2,564	2,355
12	2,636	2,412
13	2,699	2,462
14	2,755	2,507
15	2,806	2,549
16	2,852	2,585
17	2,894	2,620
18	2,932	2,651
19	2,968	2,681
20	3,001	2,709
21	3,031	2,733
22	3,060	2,758
23	3,087	2,781
24	3,112	2,802
25	3,135	2,822
Nota: para n>25 véase ISO 5725-2		

## D.2. Dócima de Cochran

1. Dado un conjunto de desviaciones típicas  $s_i$ , todas obtenidas a partir del mismo número ( $n$ ) de réplicas en la obtención de los resultados, el estadístico ( $C$ ) de la dócima de Cochran es:

$$C = \frac{s_{\max}^2}{\sum_{i=1}^p s_i^2}$$

donde:

$s_{\max}$  es la desviación típica de mayor valor dentro del conjunto

2. El resultado obtenido del estadístico se compara con su valor crítico (valor límite) para el número de muestras que interviene en el experimento ( $p$ ) y el número de réplicas ( $n$ ) dados en la tabla 2.

3. Criterios para la aceptación o rechazo.

- Si el valor estadístico ( $C$ ) es menor o igual al de su valor crítico para un  $\alpha$  del 5% se acepta como correcto.
- Si el valor estadístico es mayor al de su valor crítico para un  $\alpha$  del 5% y menor o igual al de su valor crítico para un  $\alpha$  del 1% se considera dudoso y se trata de localizar la causa asignable, si se determina se rechaza y en el caso contrario se acepta
- Si el valor del estadístico es mayor al de su valor crítico para un  $\alpha$  del 1% se considera estadísticamente incompatible y se rechaza

4. Repita la dócima de Cochran tantas veces como sea necesario.

Tabla 18. Dócima de Cochran

p	n = 2		n = 3		n = 4		n = 5		n = 6	
	1%	5%	1%	5%	1%	5%	1%	5%	1%	5%
2	-	-	0,995	0,975	0,979	0,939	0,959	0,906	0,937	0,877
3	0,993	0,937	0,942	0,871	0,883	0,798	0,834	0,746	0,793	0,707
4	0,968	0,906	0,864	0,768	0,781	0,684	0,721	0,629	0,676	0,590
5	0,928	0,841	0,788	0,684	0,696	0,598	0,633	0,544	0,588	0,506
6	0,883	0,781	0,722	0,616	0,626	0,532	0,564	0,480	0,520	0,445
7	0,838	0,727	0,664	0,561	0,568	0,480	0,508	0,431	0,466	0,397
8	0,794	0,680	0,615	0,516	0,521	0,438	0,463	0,391	0,423	0,360
9	0,754	0,638	0,573	0,478	0,481	0,403	0,425	0,358	0,387	0,329
10	0,718	0,602	0,536	0,445	0,447	0,373	0,393	0,331	0,357	0,303
11	0,684	0,570	0,504	0,417	0,418	0,348	0,366	0,308	0,332	0,281
12	0,653	0,541	0,475	0,392	0,392	0,326	0,343	0,288	0,310	0,262
13	0,624	0,515	0,450	0,371	0,369	0,307	0,322	0,271	0,291	0,243
14	0,599	0,492	0,427	0,352	0,349	0,291	0,304	0,255	0,274	0,232
15	0,575	0,471	0,407	0,335	0,332	0,276	0,288	0,242	0,259	0,220
16	0,553	0,452	0,388	0,319	0,316	0,262	0,274	0,230	0,246	0,208
17	0,532	0,434	0,372	0,305	0,301	0,250	0,261	0,219	0,234	0,190
18	0,514	0,418	0,356	0,293	0,288	0,240	0,246	0,209	0,223	0,189
19	0,496	0,403	0,343	0,281	0,276	0,230	0,238	0,200	0,214	0,181
20	0,480	0,389	0,330	0,270	0,265	0,220	0,229	0,192	0,205	0,174
21	0,465	0,377	0,318	0,261	0,255	0,212	0,220	0,185	0,197	0,167
22	0,450	0,365	0,307	0,251	0,246	0,204	0,212	0,176	0,189	0,160
23	0,437	0,354	0,297	0,243	0,238	0,194	0,240	0,172	0,182	0,155
24	0,425	0,343	0,287	0,235	0,230	0,191	0,197	0,166	0,176	0,149
25	0,413	0,334	0,278	0,228	0,222	0,185	0,190	0,160	0,170	0,144

Nota: para  $p > 25$  véase ISO 5725 – 2

Esta dócima sirve también para determinar los valores atípicos de un conjunto de recorridos, sólo se debe sustituir desviación (es) típica (s), por recorrido (s) y s por R en la fórmula.