

2005-12-22

EXACTITUD (VERACIDAD Y PRECISIÓN) DE LOS MÉTODOS Y DE LOS RESULTADOS DE MEDICIONES. GUÍA PRÁCTICA PARA EL USO DE LA NTC 3529-2 (ISO 5725-2) EN EL DISEÑO, IMPLEMENTACIÓN Y ANÁLISIS ESTADÍSTICO DE RESULTADOS



E: ACCURACY (TRUENESS AND PRECISION) OF
MEASUREMENT METHODS AND RESULTS. PRACTICAL
GUIDANCE FOR THE USE OF ISO 5725-2:1994 IN
DESIGNING, IMPLEMENTING AND STATISTICALLY
ANALYZING INTERLABORATORY REPEATABILITY AND
REPRODUCIBILITY RESULTS

CORRESPONDENCIA: esta guía es una adopción idéntica
(IDT) a la norma ISO/TR 22971:2005
Accuracy (Trueness and Precision) of
Measurement Methods and Results.
Practical Guidance for the Use of
ISO 5725-2:1994 in Designing,
Implementing and Statistically
Analyzing Interlaboratory Repeatability
and Reproducibility Results

DESCRIPTORES: métodos estadísticos – exactitud; métodos
estadísticos – precisión; métodos
estadísticos – veracidad; métodos
estadísticos – análisis de resultados.

I.C.S.: 17.020; 03.120.30

Editada por el Instituto Colombiano de Normas Técnicas y Certificación (ICONTEC)
Apartado 14237 Bogotá, D.C. - Tel. 6078888 - Fax 2221435

PRÓLOGO

El Instituto Colombiano de Normas Técnicas y Certificación, **ICONTEC**, es el organismo nacional de normalización, según el Decreto 2269 de 1993.

ICONTEC es una entidad de carácter privado, sin ánimo de lucro, cuya Misión es fundamental para brindar soporte y desarrollo al productor y protección al consumidor. Colabora con el sector gubernamental y apoya al sector privado del país, para lograr ventajas competitivas en los mercados interno y externo.

La representación de todos los sectores involucrados en el proceso de Normalización Técnica está garantizada por los Comités Técnicos y el período de Consulta Pública, este último caracterizado por la participación del público en general.

La GTC 131 fue ratificada por el Consejo Directivo del 2005-12-22.

Esta guía está sujeta a ser actualizada permanentemente con el objeto de que responda en todo momento a las necesidades y exigencias actuales.

A continuación se relacionan las empresas que colaboraron en el estudio de esta guía a través de su participación en el Comité Técnico 4 Aplicación de métodos estadísticos.

CARULLA VIVERO S.A.
COMPAÑÍA COLOMBIANA DE
CERÁMICAS S.A.–COLCERÁMICA–
COMPAÑÍA NACIONAL DE CHOCOLATES S.A.
CONSULTOR –INDEPENDIENTE, JULIO
GARCÍA SAMPEDRO

ENZIPÁN DE COLOMBIA LTDA
GAS NATURAL S.A. ESP
GLOBAL PLASTIK S.A.
INTRA-MAR SHIPPING S.A.
NUTRIANÁLISIS LTDA
SHELL COLOMBIA S.A.

Además de las anteriores, en Consulta Pública el Proyecto se puso a consideración de las siguientes empresas:

ACERÍAS DE CALDAS S.A., –ACASA–
ACERÍAS PAZ DEL RÍO S.A.
ALMACENAMIENTO Y TRANSPORTE
ESPECIALIZADO LTDA, –ALTE LTDA–
ALPINA PRODUCTOS ALIMENTICIOS S.A.
ANHÍDRIDOS Y DERIVADOS DE
COLOMBIA S.A. –ANDERCOL–
ASOCOLCAUCHOS
ASOCRETO
ATLANTIC MINERALS AND PRODUCTS
CORPORATION
ATOFINA COLOMBIA S.A.
BAVARIA S.A.
CABLES DE ENERGÍA Y DE
TELECOMUNICACIONES S.A., –CENTELSA–
CARBOQUIMICA S.A.
CEMENTOS DEL VALLE

CHALLENGER S.A.
CHICLE ADAMS S.A.
CODENSA S.A. ESP
COLOMBIANA DE AUTO PARTES S.A.
COLOMBIANA DE EXTRUSIÓN S.A., –
EXTRUCOL–
COMPAÑÍA COLOMBIANA DE TABACO S.A.,
–COLTABACO–
COMPAÑÍA DE GALLETAS NOEL S.A.
COMPAÑÍA NACIONAL DE LEVADURAS –
LEVAPÁN S.A.–
CONCONCRETO S.A.
–CORPACERO,– CORPORACIÓN DE
ACERO

–COTECMAR– (CORPORACIÓN DE
 CIENCIA Y TECNOLOGÍA PARA EL
 DESARROLLO DE LA INDUSTRIA NAVAL,
 MARÍTIMA Y FLUVIAL)
 CRISTALERÍA PELDAR S.A.
 CYGA
 ECSI S.A.
 EDITORIAL VOLUNTAD S.A.
 ELECTROMANUFACTURAS S.A.
 EMPRESA COLOMBIANA DE
 PETRÓLEOS S.A., –ECOPETROL–
 EMPRESA DE ACUEDUCTO Y
 ALCANTARILLADO DE BOGOTÁ –ESP–
 EMPRESAS PÚBLICAS DE MEDELLÍN –
 ESP–
 ESCOBAR Y MARTÍNEZ S.A.
 ESCUELA COLOMBIANA DE INGENIERÍA
 EXXON MOBIL DE COLOMBIA S.A.
 FINCA S.A.
 FRIGORÍFICO SUIZO S.A.
 FUNDACIÓN CENTRO DE CALIDAD Y
 METROLOGÍA
 –INALCEC– CORPORACIÓN INSTITUTO
 NACIONAL DE CONSULTORÍA EN
 CALIDAD
 INDEPENDIENTE – JAIRO ÁNGEL
 INDUSTRIA COLOMBIANA DE LLANTAS –
 ICOLLANTAS–
 INDUSTRIA DE ALIMENTOS ZENÚ S.A.
 INGENIO PICHICHÍ S.A.
 INSTITUTO COLOMBIANO AGROPECUARIO,
 –ICA–
 INSTITUTO COLOMBIANO DE
 PRODUCTORES DE CEMENTO, –ICPC–
 INVESA S.A.
 IVONNE BERNIER LABORATORIO LTDA.
 LARKIN LTDA.
 LHAURAVET LTDA.
 MATRICES, TROQUELES Y MOLDES CIA LTDA.
 MERCADEO DE ALIMENTOS DE
 COLOMBIA S.A. –MEALS S.A.–

METALÚRGICA CONSTRUCEL COLOMBIA S.A.,
 –METACOL–
 MINERALES INDUSTRIALES S.A.
 MOLINO EL LOBO LTDA.
 MONÓMEROS COLOMBO VENEZOLANOS –
 E.M.A.–
 NCR COLOMBIA LTDA
 NUTRIANÁLISIS LTDA
 PARABOR COLOMBIA LTDA.
 PETROQUÍMICA COLOMBIANA S.A.
 POSTOBÓN S.A.
 PRODUCTORES DE ENVASES
 FARMACÉUTICOS S.A., –PROENFAR–
 PROFICOL S.A.
 RAZA S.A.
 RENTASISTEMAS LTDA
 SCHNEIDER ELECTRIC DE COLOMBIA S.A.
 SENA CENTRO NACIONAL DE LA
 MADERA
 SENA CENTRO NACIONAL TEXTIL
 SENA REGIONAL BOGOTÁ
 SIEMENS S.A.
 SOCIEDAD DE ACUEDUCTO
 ALCANTARILLADO Y ASEO DE B/QUILLA
 –E.S.P. - TRIPLE A–
 SYNGENTA S.A.
 TECNOLOGÍA EMPRESARIAL DE
 ALIMENTOS S.A.
 THOMAS GREG & SONS DE COLOMBIA S.A.
 –IMPRESOR DE VALORES–
 UNIVERSIDAD DE ANTIOQUIA
 UNIVERSIDAD DEL VALLE
 UNIVERSIDAD JORGE TADEO LOZANO
 UNIVERSIDAD NACIONAL DE COLOMBIA
 UNIVERSIDAD NACIONAL DE COLOMBIA
 –SEDE-MEDELLÍN–
 UNIVERSIDAD NACIONAL DE
 COLOMBIA, BOGOTÁ –REVISTA
 COLOMBIANA DE ESTADÍSTICA–

ICONTEC cuenta con un Centro de Información que pone a disposición de los interesados normas internacionales, regionales y nacionales y otros documentos relacionados.

DIRECCIÓN DE NORMALIZACIÓN

INTRODUCCIÓN

La NTC 3529 (ISO 5725) consta de seis partes, cuya estructura general se ilustra en la Figura 1.

La NTC 3529-2 (ISO 5725-2) fue desarrollada como un documento de orientación para que los comités técnicos de la ISO y otras organizaciones responsables emprendieran estudios interlaboratorio para caracterizar la variabilidad de los métodos estándar de medición. En muchas disciplinas se aceptan dos medidas de variabilidad, repetibilidad y reproducibilidad, como representativas de los datos encontrados en los procesos de medición.

La repetibilidad hace referencia a la variabilidad entre mediciones hechas en muestras o materiales idénticos nominalmente, en circunstancias idénticas. Se reconoce que debido a factores desconocidos o incontrolables que influyen en el proceso de medición, usualmente las mediciones repetidas no coincidirán. El grado de esta variabilidad se puede expresar por una desviación estándar denominada desviación estándar de repetibilidad, de los resultados de comparaciones dentro de laboratorio.

La reproducibilidad hace referencia a la variabilidad entre mediciones hechas en muestras o materiales idénticos en condiciones diferentes por diferentes laboratorios que utilizan el mismo método de medición normalizado. La reproducibilidad incluye los efectos causados por diferencias entre instrumentos, reactivos, operadores, laboratorios y condiciones ambientales. La variabilidad de resultados en estas condiciones se puede describir mediante una desviación estándar denominada desviación estándar de reproducibilidad.

Este documento de guía se divide en cuatro partes, además del alcance (numeral 1):

- Numeral 2. La organización de un programa interlaboratorio tiene que ver con la organización de la prueba interlaboratorio y comprende las funciones del coordinador de laboratorio, el personal de laboratorio y un estadístico, para la preparación y administración de la prueba; la selección de materiales y niveles de interés para la prueba; y la selección de laboratorios. También describe cómo el número de réplicas de mediciones (que se van a hacer en cada muestra) se debe tratar estadísticamente y la forma en que los datos resultantes se deben reportar.
- Numeral 3. Examen crítico de los datos; trata sobre los datos que usan procedimientos gráficos y numéricos. Se da orientación relativa a cuando los datos son anómalos, es decir, si son inconsistentes con otros datos del estudio, y para pruebas de valores atípicos que se usan para identificar la presencia o ausencia de datos anómalos.
- Numeral 4. Estimación de las desviaciones estándar de repetibilidad y reproducibilidad; trata sobre la estimación e interpretación de la repetibilidad de las desviaciones estándar de repetibilidad y reproducibilidad. También se incluye una comparación de las contribuciones relativas de las desviaciones estándar de repetibilidad y reproducibilidad.

- Numeral 5, ejemplos aplicados en los que se utiliza software estadístico; trata sobre ejemplos que resaltan las diferentes técnicas que se pueden usar.

Se recomienda que este documento guía se lea conjuntamente con la NTC 3529-2 (ISO 5725-2) y no se debería usar en reemplazo de dicha norma.

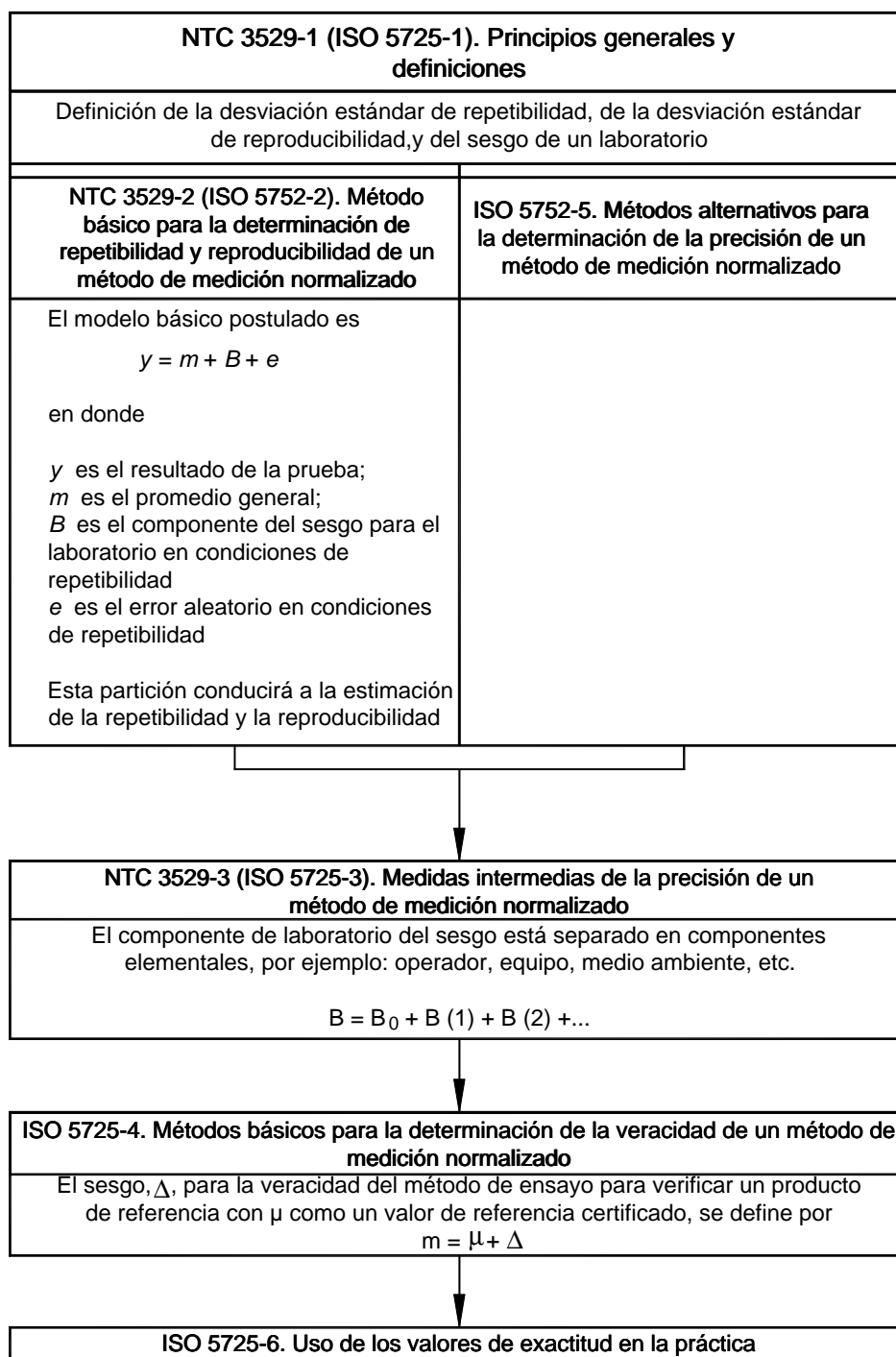


Figura 1. Estructura de la ISO 5725. Aplicación de un método de ensayo normalizado para los análisis de una muestra o producto en diferentes laboratorios

EXACTITUD (VERACIDAD Y PRECISIÓN) DE LOS MÉTODOS Y DE LOS RESULTADOS DE MEDICIONES. GUÍA PRÁCTICA PARA EL USO DE LA NTC 3529-2 (ISO 5725-2) EN EL DISEÑO, IMPLEMENTACIÓN Y ANÁLISIS ESTADÍSTICO DE RESULTADOS DE REPETIBILIDAD Y REPRODUCIBILIDAD INTERLABORATORIOS

1. OBJETO Y ALCANCE

Este documento proporciona orientación práctica para el uso de la NTC 3529-2 (ISO 5725-2), y presenta procedimientos graduales simplificados para el diseño, implementación y análisis estadísticos de estudios interlaboratorio para evaluar la variabilidad de un método de medición normalizado, y sobre la determinación de la repetibilidad y reproducibilidad de los datos obtenidos en pruebas interlaboratorio.

2. ORGANIZACIÓN DE UN PROGRAMA INTERLABORATORIO

2.1 REQUISITOS DE UN EXPERIMENTO DE PRECISIÓN

El experimento en su totalidad está organizado para:

- a) Suministrar un conjunto completo de resultados de:
 - Número p_{Lab} de laboratorios que demuestran que el procedimiento de ensayo está bien controlado, y para cuantificar la dispersión observada, estimada por la reproducibilidad.
- NOTA El símbolo p_{Lab} usado en este documento tiene el mismo significado que el símbolo p usado en la NTC 3529-2 (ISO 5725-2). El cambio se hizo para diferenciar claramente este símbolo del símbolo P usado para "probabilidad". Algunas veces es difícil diferenciar las P mayúscula y minúscula, particularmente cuando son subíndices.
- Número q de muestras o productos que representan diferentes niveles de resultados o desempeño. Un valor mínimo para q es dos, pero valores de cinco a diez son más apropiados para demostrar que el procedimiento de ensayo está en capacidad de discriminar correctamente entre niveles.
 - Número n de celdas de réplicas que demuestran que el procedimiento de ensayo está bien controlado dentro de un solo laboratorio. Cuando el número de laboratorios y de niveles es suficiente, se requieren al menos dos determinaciones.

- b) Analizar estadísticamente (véanse los numerales 2 y 3) una tabla de resultados reportados por p_{Lab} que analizan q muestras, ensayadas n veces en condiciones de repetibilidad.

La tabla de los resultados presentados al coordinador de laboratorio se ilustra en la Tabla 1 (véase la NTC 3529-2 (ISO 5725-2), numeral 7.2.8).

2.2 RESPONSABILIDADES DEL PERSONAL INVOLUCRADO EN UN EXPERIMENTO DE PRECISIÓN

2.2.1 Generalidades

Un programa interlaboratorio es muy costoso, tanto en términos de su coordinación como de su participación. En consecuencia, el ensayo de desempeño debería estar bien coordinado y planificado. En cualquier programa interlaboratorio es necesario considerar tres tipos de actividades, como se ilustra en la Figura 2.

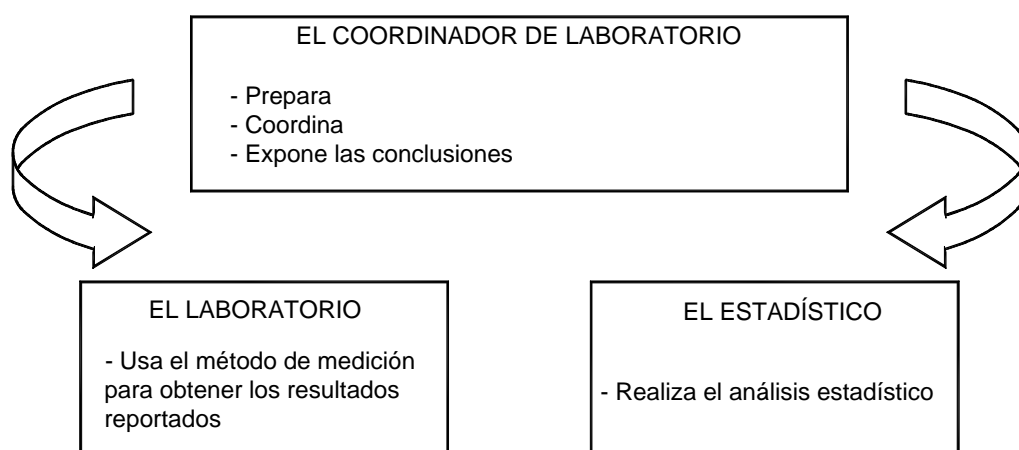


Figura 2. Responsabilidades de las funciones ejecutivas

2.2.2 Coordinador de laboratorio

Las principales tareas del coordinador de laboratorio son:

- Organizar el programa interlaboratorio, con la asesoría del estadístico para la elaboración del diseño experimental.
- Coordinar el progreso.
- Exponer las conclusiones.

Los deberes del coordinador de laboratorio los puede realizar más de una persona. Sin embargo, una sola persona debería ser responsable de todo el programa. El coordinador de laboratorio debería estar familiarizado con el método estándar, pero no debería participar en el proceso de medición.

2.2.3 Laboratorio

El personal de laboratorio debería tener considerable experiencia con el método de medición de ensayo.

El laboratorio debe emprender el análisis siguiendo los procedimientos de ensayo recibidos del coordinador de laboratorio. Cualquier comentario emitido por el laboratorio sobre el uso del método de ensayo se debería reportar al coordinador de laboratorio. Sin embargo, los procedimientos realizados por el laboratorio deberían ser los que suministre dicho coordinador.

El laboratorio debe cumplir los requisitos indicados por el coordinador de laboratorio, que incluyen:

- Almacenamiento de muestras.
- Fecha y orden de realización del análisis.

El laboratorio debe suministrar al coordinador de laboratorio los resultados del análisis de la manera establecida por el coordinador de laboratorio.

2.2.4 Estadístico

El estadístico debe recibir del coordinador de laboratorio los datos sin procesar obtenidos usando el método indicado como se definió en la Tabla 1.

El estadístico debe examinar los datos y aplicar cualquier prueba estadística, de preferencia las pruebas descritas en la NTC 3529-2 (ISO 5725-2), para identificar valores atípicos potenciales. Los valores atípicos se deben enseñar al coordinador de laboratorio, quien debe llevar a cabo la investigación apropiada para determinar cuándo retener, rechazar o modificar los datos.

El estadístico debe llevar a cabo los análisis estadísticos, elaborar los gráficos y brindar estimados de las medias y varianzas (NTC 3529-2 (ISO 5725-2), 7.1.2). El estadístico debe resumir todos los resultados de los análisis estadísticos en un informe que se debe enviar al coordinador de laboratorio.

3. EXAMEN CRÍTICO DE LOS DATOS

3.1 DESCRIPCIÓN DE LOS DATOS

3.1.1 Datos sin procesar

Los datos se presentan como se ilustra en la Tabla 1. Las Tablas 2 y 3 se elaboran a partir de la Tabla 1. Algunos paquetes de software estadístico validados pueden ofrecer diferentes presentaciones de la misma información.

Tabla 1. Comparación de los datos sin procesar

Laboratorio	Nivel						
	1	2	...	j	...	$q - 1$	q
1							
2							
...							
i				y_{ij1} y_{ij2} y_{ijn}			
...							
P_{lab}							

Tabla 2. Comparación de los valores medios para cada celda de la Tabla 1

Laboratorio	Nivel						
	1	2	...	j	...	$q - 1$	q
1							
2							
...							
i				m_{ij}			
...							
P_{lab}							

Tabla 3. Comparación de valores que indican la dispersión^a de los valores para cada celda en la Tabla 1

Laboratorio	Nivel						
	1	2	...	j	...	$q - 1$	q
1							
2							
...							
i				s_{ij}			
...							
P_{lab}							

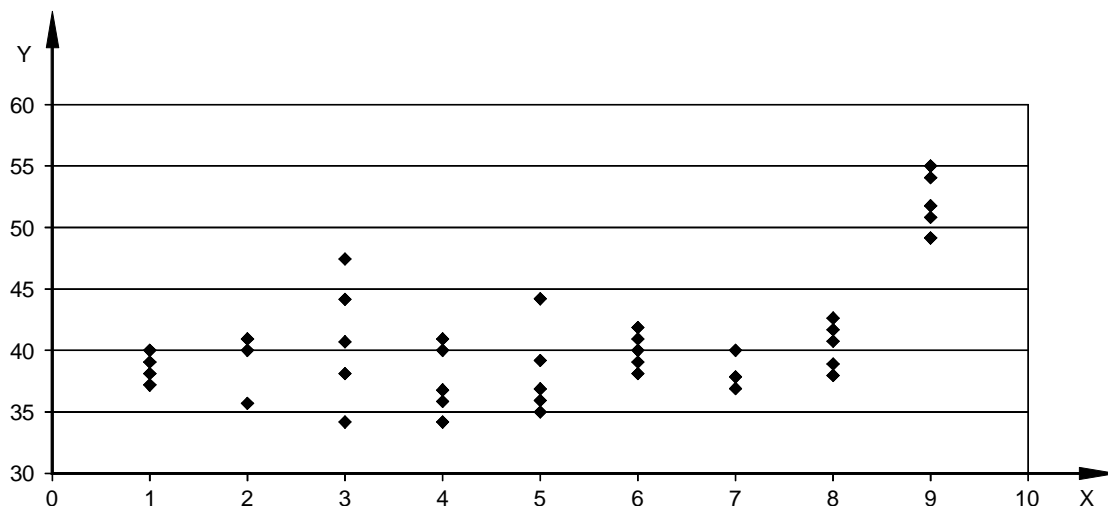
^a La medida más común de dispersión es la desviación estándar.

3.1.2 Representación gráfica de los datos

3.1.2.1 Gráfico de resultados por laboratorios (gráfico de datos sin procesar)

Antes de realizar cualquier ensayo para determinar valores atípicos potenciales, se recomienda elaborar gráficos de los datos sin procesar. De esta forma es posible obtener una imagen “instantánea” de los resultados, por ejemplo, como se ilustra en la Figura 3 (que está basada en la NTC 3529-2 (ISO 5725-2), Figuras B.1 a B.4). Se puede obtener gran cantidad de información mediante la inspección visual de un gráfico de los datos sin procesar, y una evaluación instantánea de la dispersión de datos verificados. En consecuencia, se puede sugerir una indicación de la presencia de valores atípicos, o podrían llegar a ser evidentes diferencias inusuales a niveles de interés particular, simplemente mediante la inspección visual de la gráfica de datos apropiada. Por ejemplo, en la Figura 3, el gráfico de resultados para el laboratorio 3 podría sugerir una dispersión mayor de la esperada de resultados comparados con todos los demás laboratorios; en consecuencia, la repetibilidad total se verá afectada. Esta posibilidad se puede confirmar mediante la prueba de *Cochran*. Además, los resultados para el

laboratorio 9 podrían sugerir un valor atípico respecto al valor medio del laboratorio cuando se compara con los otros valores medios para todos los laboratorios de ensayo. En consecuencia, la reproducibilidad se podría ver afectada y esto se puede confirmar mediante la prueba de *Grubbs*.



Convenciones:

X Identificación del laboratorio
Y Resultados por laboratorio

Figura 3. Representación gráfica de datos sin procesar para un nivel de interés particular

3.1.2.2 Diagrama de caja y bigotes

En donde se reportan muchos resultados, especialmente para un nivel de interés particular, un diagrama de caja y bigotes puede revelar información similar a la del numeral 3.1.2.1; en la Figura 4 se presenta un ejemplo. Sin embargo, este tipo de gráfico, que se basa en estadísticos robustos que incluyen la determinación del valor de la mediana, no se describe en la NTC 3529-2 (ISO 5725-2). No obstante, se define e ilustra en los ejemplos del numeral 4, ya que estos gráficos se encuentran disponibles en la mayoría de software estadístico.

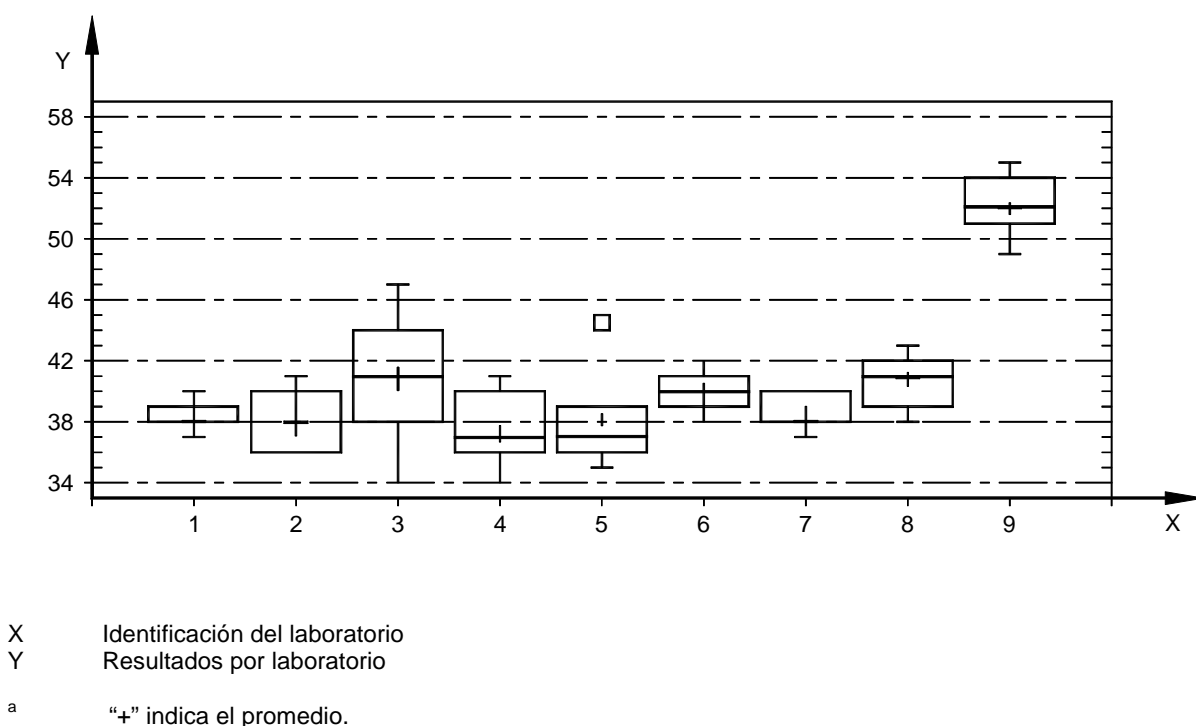
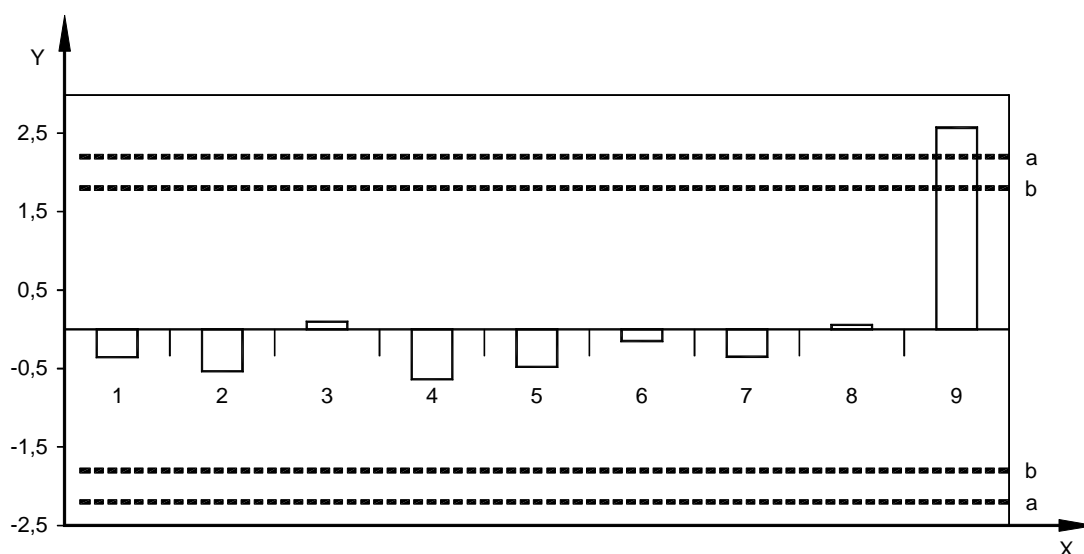


Figura 4. Gráfico de caja y bigotes

3.1.2.3 Gráficos de Mandel de estadísticos de pruebas h y k

3.1.2.3.1 Gráfico h de Mandel

Para un nivel de interés particular, los valores medios obtenidos para todos los laboratorios se usan con el propósito de calcular un solo valor medio total. Este valor se usa entonces para calcular el estadístico h de Mandel para todos los laboratorios para este nivel. Este estadístico se define en la NTC 3529-2 (ISO 5725-2), ecuación (6). Este estadístico es la relación de la diferencia entre la media para un conjunto particular de datos, y la media de todos los conjuntos de datos, y la desviación estándar de las medias de todos los conjuntos de datos. Este cociente se grafica y se compara con los valores de la relación calculados o tabulados obtenidos para los niveles de confianza del 95 % y del 99 %. El mismo procedimiento se utiliza para calcular el estadístico h de *Mandel* para todos los laboratorios, para todos los otros niveles de interés (véase la Figura 5). Vale la pena señalar que se pueden trazar valores tanto positivos como negativos.



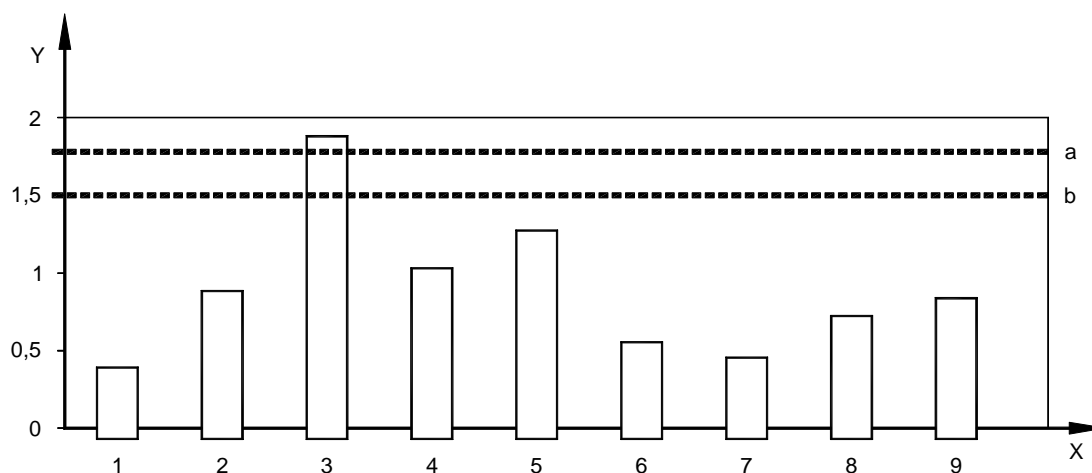
Convenciones

X	Identificación del laboratorio
Y	<i>h</i> de <i>Mandel</i>
a	nivel de confianza del 99 %
b	nivel de confianza del 95 %

Figura 5. Gráfico *h* de *Mandel*

3.1.2.3.2 Gráfico *k* de *Mandel*

Para un nivel de interés particular, las desviaciones estándar obtenidas para todos los laboratorios se usan para calcular una desviación media estándar o una desviación estándar sencilla agrupada. Este valor se usa para calcular el estadístico *k* de *Mandel* para todos los laboratorios de este nivel. Este estadístico se define en la NTC 3529-2 (ISO 5725-2), ecuación (7). Es el cociente de la desviación estándar de los resultados y la desviación estándar media o agrupada. Este cociente se grafica y se compara con los valores de las relaciones calculadas o tabuladas obtenidas para los niveles de confianza del 95 % y del 99 %. El mismo procedimiento se usa entonces para calcular el estadístico *k* de *Mandel* para todos los laboratorios, para todos los otros niveles de interés (véase la Figura 6). Vale la pena señalar que solamente se grafican los valores positivos.



Convenciones

X Identificación del laboratorio
Y k de Mandel

a nivel de confianza del 99 %
b nivel de confianza del 95 %

Figura 6. Gráfico k de Mandel

3.1.2.3.3 Inspección gráfica

De los gráficos, se pueden identificar resultados individuales para cada laboratorio, que podrían considerarse diferentes de la distribución esperada de resultados. Por ejemplo, el gráfico h para niveles particulares de interés para cada laboratorio podría aproximarse o exceder el valor estadístico h de Mandel calculado al nivel de confianza del 95 % ó 99 %, si la prueba de *Grubbs* indica presencia de valores atípicos. Además, el gráfico k para niveles de interés particulares para cada laboratorio podría aproximarse o exceder el valor estadístico k calculado al nivel de confianza del 95 % ó 99 % si la prueba de *Cochran* indica presencia de valores atípicos.

3.2 PRUEBAS PARA VALORES ATÍPICOS

3.2.1 Puntos generales

3.2.1.1 Nivel de confianza

El tratamiento de los valores atípicos se aborda en la NTC 3529-2 (ISO 5725-2), numeral 7, particularmente del 7.1 al 7.3. Un valor atípico se puede considerar como un resultado que es suficientemente diferente de todos los otros resultados, como para justificar investigación adicional. Dependiendo del tipo de distribución a la que se ajustan los resultados, un resultado que parece ser un valor atípico podría en realidad ser un resultado válido. La NTC 3529-2 (ISO 5725-2), numerales 7.3.2.1 y 7.3.3.2 recomienda niveles de confianza del 95 % para valores atípicos denominados anormales, y el 99 % para valores atípicos denominados “valores atípicos estadísticos”. Para circunstancias individuales, la selección de niveles de confianza del 95 % y 99 % significa que un resultado en 20 y un resultado en 100, respectivamente, se podrían interpretar erróneamente. En consecuencia, este resultado podría ocurrir por casualidad y el grado de confianza indicado en la NTC 3529-2 (ISO 5725-2) podría no ser apropiado para las necesidades individuales. Esto podría representar un grado de aceptabilidad que no es suficiente para algunos propósitos. Esto significaría que las

circunstancias individuales ameritarían considerarse individualmente en cuanto a si se debería aplicar la NTC 3529-2 (ISO 5725-2), en términos de los niveles de confianza usados.

3.2.1.2 Suposiciones básicas

En los ensayos usados para determinar la presencia o ausencia de valores atípicos, se supone que los resultados se distribuyen de una manera *Gaussiana* (denominada comúnmente como distribución normal; la NTC 3529-2 (ISO 5725-2), numeral 1.4) o al menos en una distribución unimodal sencilla (NTC 3529-2 (ISO 5725-2), numeral 7.3.1.7). En consecuencia, antes de emprender cualquier ensayo, especialmente uno que involucre un gran número de resultados, es conveniente llevar a cabo una revisión para confirmar esta suposición. También se supone (NTC 3529-2 (ISO 5725-2), numerales 1.3 y 5.1.1) que el número de resultados dentro de cada conjunto de datos (de cada laboratorio) es el mismo y que el número de resultados para cada nivel de interés, o el número de muestras diferentes, es el mismo. Así, los resultados están “equilibrados”. Si los resultados no están “equilibrados”, entonces se recomienda (NTC 3529-2 (ISO 5725-2), numeral 7.2.2) que los resultados de conjuntos apropiados de datos se descarten aleatoriamente hasta que se cree una situación “equilibrada”. Aunque se prefiere una situación “equilibrada”, se reconoce (incluso dentro de los ejemplos ilustrados en la NTC 3529-2 (ISO 5725-2)) que se puede dar cabida a situaciones “no equilibradas”. Además, se supone que los resultados se obtienen en condiciones de repetibilidad (NTC 3529-1 (ISO 5725-1:1994), numeral 4.4 y NTC 3529-2 (ISO 5725-2), numeral 7.3.3.3). En consecuencia, se puede suponer que las muestras para un nivel de interés específico son homogéneas, idénticas en todos los aspectos, y que se analizan en un período de tiempo corto usando los mismos reactivos y soluciones de calibración. En teoría, estos criterios se tienen que satisfacer antes de que se pueda aplicar cualquier ensayo para determinar la presencia o ausencia de valores atípicos.

3.2.1.3 Declaración de valores atípicos

Cuando se llevan a cabo ensayos de valores atípicos, se debe comprender que los valores atípicos no se deberían descartar ni rechazar solamente desde un punto de vista estadístico. Para cada muestra, se debería investigar e identificar la razón por la cual el resultado es diferente de todos los otros. Los ensayos de valores atípicos (basados en las hipótesis usadas) indican si existe una causa estadística suficiente para un valor atípico; no indicará por qué ha ocurrido. Solamente después de profundas investigaciones que se lleven a cabo para determinar las causas probables, es que los datos se deberían declarar como valores atípicos y desecharlos.

Cuando se ha analizado un nivel de interés particular mediante la prueba de *Cochran*, de *Grubbs* o cualquier otra y no se hayan identificado valores atípicos o valores atípicos adicionales, entonces se ensayan en forma similar otros niveles de interés. Si se identifican varios valores atípicos en diferentes niveles de interés para un conjunto dado de datos producidos por un solo laboratorio, puede ser necesario considerar si todos los conjuntos de datos para todos los niveles de interés se deberían investigar adicionalmente.

Además, se debería considerar desechar solamente los valores atípicos identificados para ese laboratorio, o rechazar el conjunto entero de datos para ese laboratorio. La experiencia en esta materia determinará el curso de acción necesario y se debería realizar en forma individual y con conocimiento de las investigaciones realizadas para identificar las posibles causas.

La inspección de los gráficos de valores atípicos puede brindar evidencia adicional a la que suministraron los ensayos de valores (numéricos) atípicos. En consecuencia, los gráficos *h* y *k* de *Mandel* se pueden usar para facilitar estas decisiones. Se debería considerar rechazar todos los resultados de un laboratorio si un conjunto particular de datos (por ejemplo, para el

laboratorio 9 en la Figura 5 ó para el laboratorio 3 en la Figura 6) muestra que todos los valores calculados son positivos y se aproximan o exceden los valores tabulados a los niveles de confianza del 95 % y el 99 %. Podría ser que en el ejemplo escogido todos los valores medios para este laboratorio sean mayores que los valores correspondientes para todos los otros laboratorios. Este hecho puede ser motivo de preocupación. Como antes, la decisión de rechazar o descartar datos se debe tomar solamente a la luz de investigaciones apropiadas realizadas para determinar la causa probable de valores atípicos.

3.2.1.4 Selección de las pruebas

Para cada laboratorio o nivel de interés o muestra particular, la mayoría de pruebas de valores atípicos comparan alguna medida de la distancia relativa de un resultado sospechoso, con la media de todos los resultados de ensayo, y evalúan esta comparación para determinar si el resultado pudo ocurrir por casualidad. Hay muchas pruebas disponibles, pero para propósitos prácticos, no todas se describen completamente en la NTC 3529-2 (ISO 5725-2), como se indica en la NTC 3529-2 (ISO 5725-2), numeral 7.1.3. Así, además de las pruebas mencionadas en la NTC 3529-2 (ISO 5725-2), numerales del 7.1 al 7.3, se debe entender que también se pueden usar otras pruebas para determinar valores atípicos potenciales. En consecuencia, el uso de pruebas de valores atípicos que no se describen en la NTC 3529-2 (ISO 5725-2) se deja a discreción del estadístico. Los detalles completos de la prueba se deben registrar con los resultados.

3.2.2 Prueba de Cochran

3.2.2.1 Principio

La prueba de *Cochran* se usa para verificar si hay desviaciones estándar para celdas en la Tabla 3, que son excepcionalmente grandes e inflarían el estimado de desviación estándar de repetibilidad si se conservan. El estadístico usado en la prueba de *Cochran* está relacionado estrechamente con el estadístico k de *Mandel*.

3.2.2.2 Interpretación

De la Tabla 3, la prueba de *Cochran* se puede usar para identificar si la mayor variación de un conjunto de datos dado es compatible con las variaciones en los datos de todos los conjuntos de datos para ese nivel de interés. Esta prueba posibilita calcular el cociente de la varianza más grande obtenida para un laboratorio particular, y la suma de todas las varianzas obtenidas para todos los laboratorios. Este valor calculado se compara entonces con la relación calculada o tabulada (crítica) y se evalúa la presencia de valores atípicos anormales o valores atípicos.

La prueba de *Cochran* identificará aquellas varianzas que son mayores que las varianzas esperadas para ese nivel de interés. A este respecto, la prueba es una prueba unilateral, ya que el laboratorio con la menor varianza (en relación con otros laboratorios) no se somete a esta prueba. La decisión de repetir la prueba de *Cochran* dependerá de si se identifica un valor atípico, y del número del conjunto de datos que se van a ensayar para ese nivel de interés particular. Si no se indica un valor atípico, entonces la prueba no se repite. Si se indica un valor atípico, se debe tener la precaución de repetir la prueba de *Cochran* en los conjuntos de datos restantes para ese nivel de interés. Esta precaución es particularmente pertinente para pequeñas cantidades de datos, especialmente si aproximadamente el 20 % de los datos se rechazan finalmente como valores atípicos.

3.2.2.3 Pruebas alternativas

La prueba de *Bartlett*, la de *Levene* y la de *Hartley* son alternativas a la de *Cochran*. Sin embargo, puede haber ocasiones, particularmente en los casos dudosos, en que se identifican valores atípicos usando una prueba, pero no se identifican usando otra prueba. En consecuencia, es importante que el estadístico reporte qué ensayos se han usado y las conclusiones alcanzadas.

3.2.3 Prueba de *Grubb*

3.2.3.1 Principio

La prueba de *Grubb* se usa para verificar si hay medias de celda en la Tabla 2, que son excepcionalmente altas o bajas e inflarían el estimado de la desviación estándar de reproducibilidad si se mantienen. El estadístico usado en la prueba de *Grubb* está estrechamente relacionado con el estadístico h de *Mandel*.

3.2.3.2 Interpretación

Después de que se ha realizado la prueba de *Cochran*, los valores medios tabulados para cada nivel de interés particular presentado en la Tabla 2 se disponen en orden no decreciente. Se realizan varias pruebas de *Grubb*. Primero, la prueba se realiza para determinar si el valor medio más alto o más bajo se puede identificar como un valor atípico único. Si se detecta un valor atípico, se descarta y la prueba se repite para el otro valor extremo. Para un nivel de interés particular, un ensayo de *Grubbs* para un valor atípico permite el cálculo del cociente de la diferencia entre el valor sospechoso y la media de todos los valores para ese nivel, y la desviación estándar de todos los valores. Esta relación se compara entonces con los valores de la relación calculada o tabulada (crítica) a los niveles de confianza del 95 % y el 99 %.

Si no se identifica un valor atípico único, se realiza una prueba de *Grubb* adicional para determinar la presencia (o ausencia) de dos valores atípicos extremos. Por ejemplo, los dos valores medios más bajos se ponen a prueba, y si no se observa ningún valor atípico, se ponen a prueba los valores medios más altos. En esta prueba, si el cociente calculado es más que la relación calculada al nivel de confianza establecido, las medias se pueden considerar satisfactorias.

3.2.3.3 Pruebas alternativas

Una alternativa a la prueba de *Grubb* es la prueba de *Dixon*. Sin embargo, de nuevo habrá ocasiones, particularmente en los casos dudosos, en las que se identifica un valor atípico usando una prueba, pero no se identifica usando otra. En consecuencia, es importante que el estadístico reporte qué pruebas se han usado.

3.3 CONCLUSIONES

El diagrama de flujo de la Figura 7 destaca los procedimientos principales que se deben llevar a cabo.

Los datos que se han identificado como valores atípicos estadísticos se reportan al coordinador de laboratorio. El coordinador debe llevar a cabo una investigación apropiada para determinar si retener, rechazar o modificar cualquier dato. Cuando esta investigación haya finalizado y dependiendo de su resultado, el estadístico puede recibir un conjunto de datos revisados en el mismo formato de la Tabla 1. Si es apropiado, esto permite recalcular tablas adicionales similares a las Tablas 2 y 3.

El estadístico entonces puede estimar para cada nivel de interés las desviaciones estándar de repetibilidad y reproducibilidad.

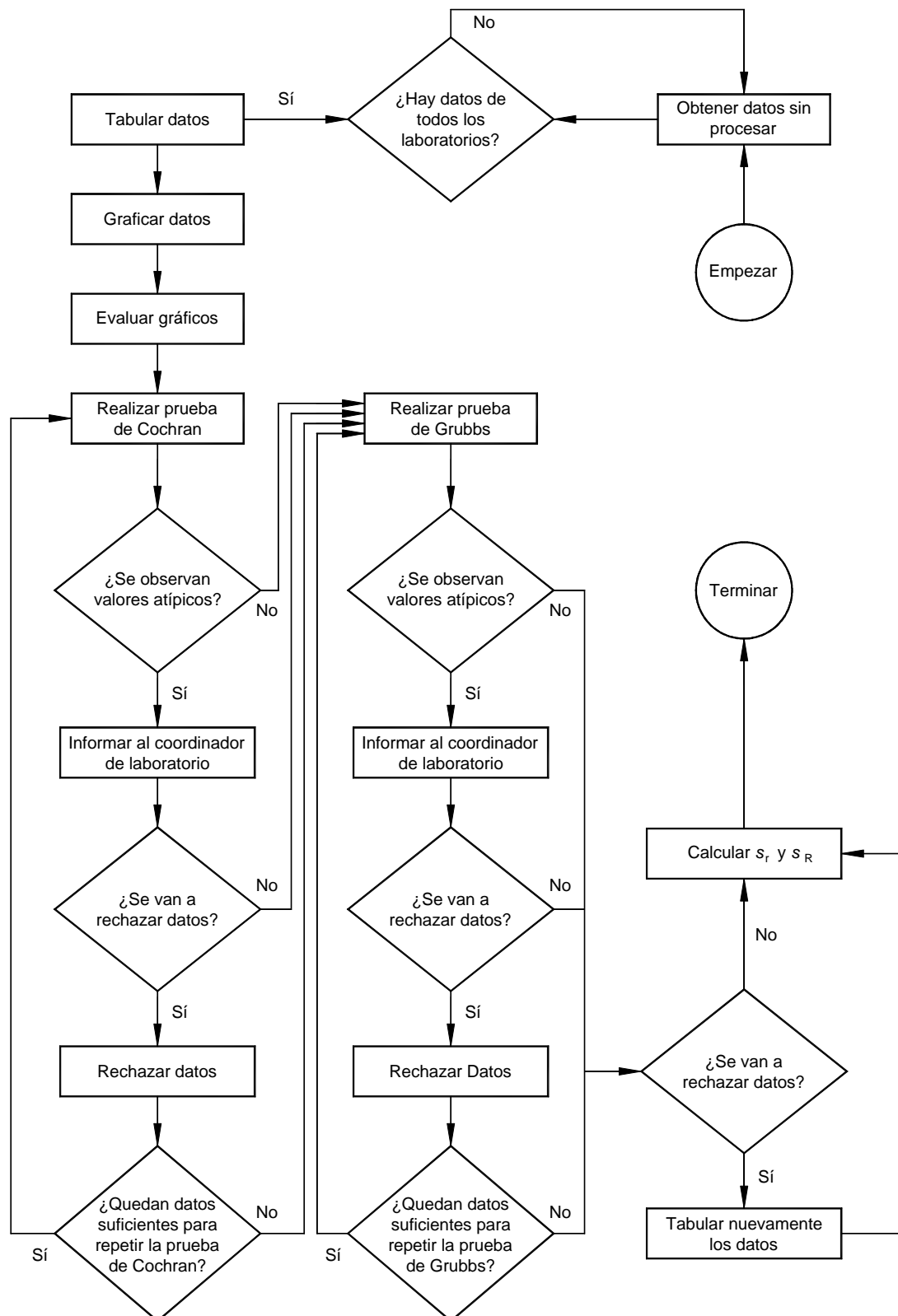


Figura 7. Diagrama de flujo para el tratamiento estadístico de valores atípicos

4. ESTIMACIÓN DE LAS DESVIACIONES ESTÁNDAR DE REPETIBILIDAD Y REPRODUCIBILIDAD

4.1 ANÁLISIS DE VARIANZA

En un programa interlaboratorio, un gran número de laboratorios lleva a cabo pruebas repetidas sobre la misma muestra de material. El diseño es como el que se ilustra en la Figura 8.

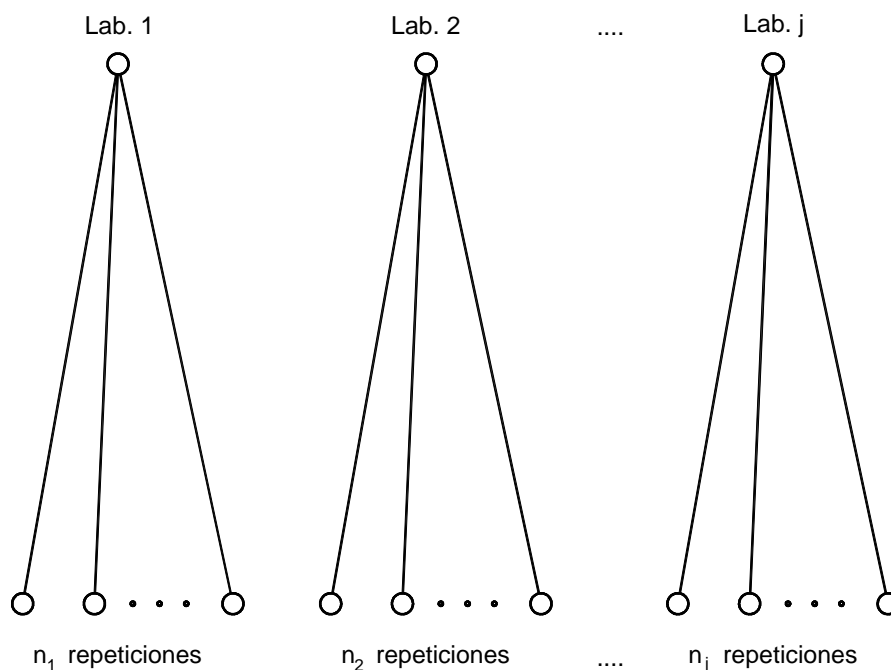


Figura 8. Diseño del programa interlaboratorio

Si la muestra se ensaya una vez en diferentes laboratorios, la variación de la característica resultante y reflejará alguna combinación de la variabilidad que surge de la variación en laboratorio y de la variabilidad que surge de la variación entre laboratorios. El modelo descrito en el numeral 4.2 se usa para explicar la generación de los resultados individuales, y .

4.2 DESCRIPCIÓN DEL MODELO

La precisión (repetibilidad y reproducibilidad) de un método de medición se puede estimar del análisis de datos de un grupo de laboratorios seleccionados de una población de laboratorios usando el mismo método. Cada laboratorio brinda un conjunto de resultados que se compara estadísticamente con otro conjunto de resultados.

NOTA Estos procedimientos también se pueden usar para estimar los parámetros de un grupo de analistas u operadores, en lugar de un grupo de laboratorios.

El tratamiento estadístico se basa en lo siguiente:

- Varias muestras a diferentes niveles de resultados o desempeño, son analizadas por varios laboratorios.
- Cada laboratorio lleva a cabo el método de prueba y reporta los resultados por duplicado.

- A cada nivel, la dispersión total de los resultados se desglosa en variaciones aleatorias.

Este proceso se define en la NTC 3529-2 (ISO 5725-2), numeral 4.1, por la ecuación:

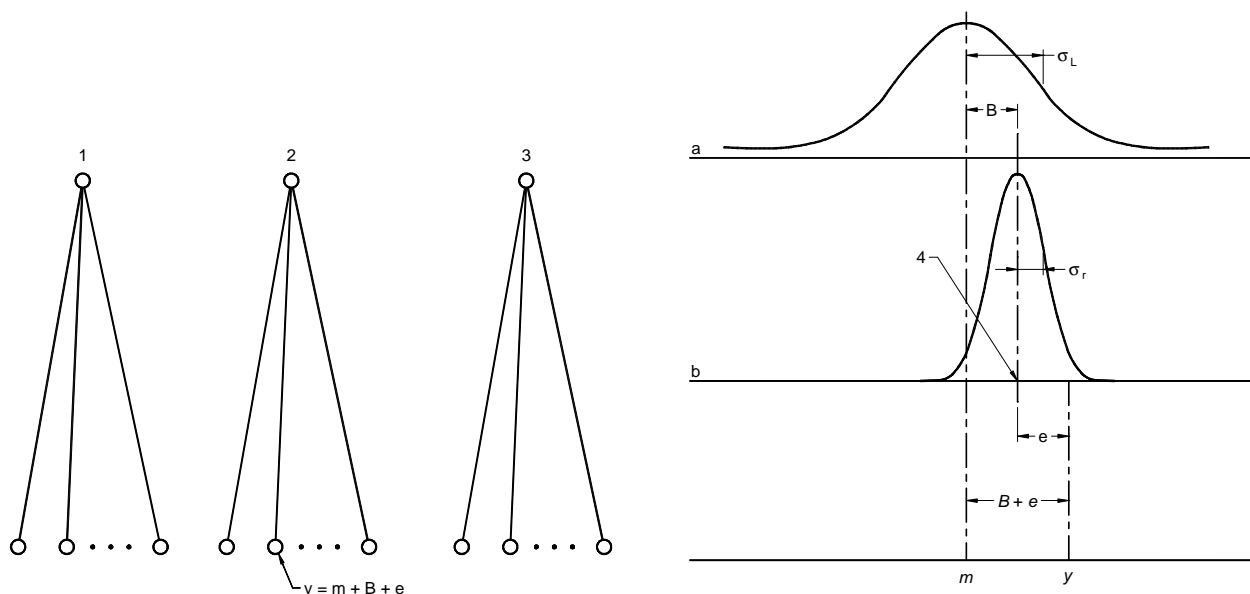
$$y = m + B + e$$

en donde

- m es la media de los resultados;
- B es el componente de laboratorio del sesgo en condiciones de repetibilidad;
- e es la variabilidad aleatoria que ocurre durante cualquier medición, en condiciones de repetibilidad.

El mejor estimado disponible del valor real de la característica que se estudia es m , la media total de todos los resultados. En general, un valor reportado único, y , no será igual a m .

El error total ($y - m$) contiene dos partes: e y B , en donde e es la desviación de y de la media de un gran número de resultados que pueden provenir de un laboratorio particular, en donde se generó y , y B es la desviación (o diferencia) entre esta media conceptual y la media total. Esta partición, $y = m + B + e$ se ilustra en la Figura 9.



Convenciones:

1. Lab. 1
2. Lab. 2
3. Lab. j
4. media del Lab. 2

- a Variación del componente laboratorio del sesgo.
- b Variación de los resultados de las pruebas dentro de un laboratorio

Figura 9. Partición del error total

Se espera que e y B estén cerca de cero y el propósito de la NTC 3529-2 (ISO 5725-2) es especificar cómo estimar sus varianzas. La varianza de e es la varianza en laboratorio, $s^2_{e_L}$, y la varianza de B es la varianza entre laboratorios, $s^2_{B_L}$.

La varianza en laboratorio se conoce como la varianza de la repetibilidad.

La suma de la varianza entre laboratorios y la varianza en laboratorio $s^2_R = s^2_{B_L} + s^2_{e_L}$ representa la varianza de los resultados obtenidos en condiciones que son apreciablemente diferentes (tales como un cambio de laboratorios) y se conoce como varianza de la reproducibilidad, s^2_R .

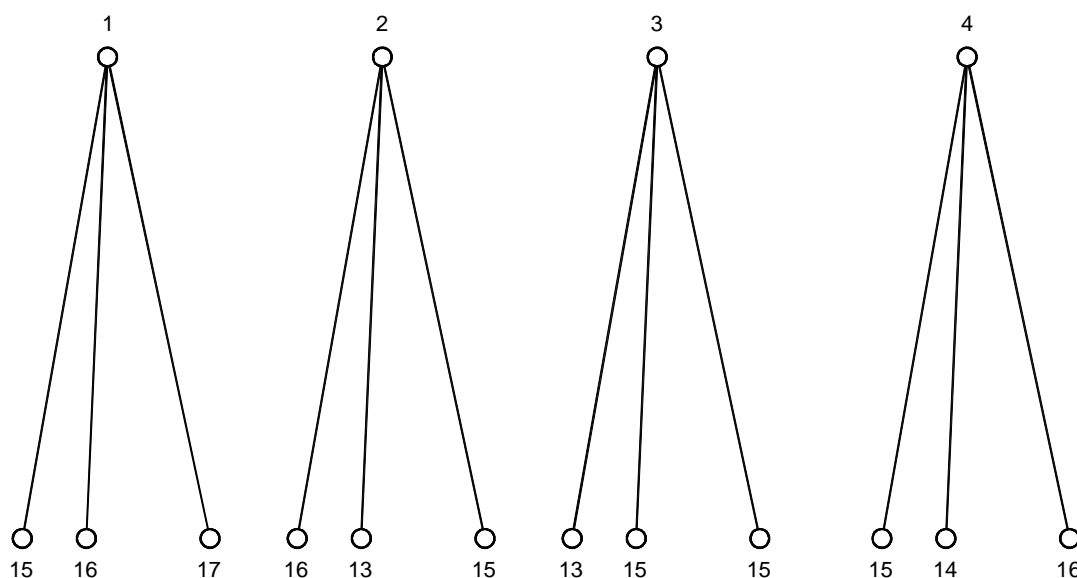
La estructura descrita en la Figura 8 se denomina diseño anidado o jerárquico. Por tanto, es necesario estimar $s^2_{e_L}$ y $s^2_{B_L}$ para evaluar la varianza de e y B .

Estadísticamente hablando, se dice que el efecto del laboratorio, B , es aleatorio. Esto significa que el conjunto de laboratorios que participaron en el programa interlaboratorio es una selección aleatoria de un infinito conceptual, por el número de laboratorios que podría usar este método de ensayo. Es necesario estimar la compatibilidad entre estos laboratorios entre sí.

4.3 EJEMPLOS

4.3.1 Ejemplo 1

Para explicar cómo se pueden estimar los componentes de la varianza, $s^2_{e_L}$ y $s^2_{B_L}$, suponga que hay cuatro laboratorios participantes, cada uno de los cuales suministra tres repeticiones de la misma muestra, sin datos faltantes ni valores atípicos. Por ejemplo, los datos reportados son como se ilustran en la Figura 10.



Convenciones

1. Lab 1
2. Lab. 2
3. Lab 3
4. Lab 4

Figura 10. Diseño del ejemplo 1

En el Lab 1, un estimado $s^2_{r(Lab\ 1)}$ de s^2_r está dado por la varianza de los tres resultados, en consecuencia:

$$s^2_{r(Lab\ 1)} = \text{var} (15, 16, 17) = 1,00 \text{ con dos grados de libertad.}$$

En el Lab 2, un estimado $s^2_{r(Lab\ 2)}$ de s^2_r está dado por la varianza de los tres resultados, en consecuencia:

$$s^2_{r(Lab\ 2)} = \text{var} (16, 13, 15) = 2,33 \text{ con dos grados de libertad.}$$

En Lab 3, un estimado $s^2_{r(Lab\ 3)}$ de s^2_r está dado por la varianza de los tres resultados, en consecuencia:

$$s^2_{r(Lab\ 3)} = \text{var} (13, 15, 15) = 1,33 \text{ con dos grados de libertad.}$$

En Lab 4, un estimado $s^2_{r(Lab\ 4)}$ de s^2_r está dado por la varianza de los tres resultados, en consecuencia:

$$s^2_{r(Lab\ 4)} = \text{var} (15, 14, 16) = 1,00 \text{ con dos grados de libertad.}$$

Un mejor estimado, s^2_r , de la varianza de repetibilidad desconocida real, s^2_r se obtiene promediando $s^2_{r(lab\ 1)}$, $s^2_{r(lab\ 2)}$, $s^2_{r(lab\ 3)}$ y $s^2_{r(lab\ 4)}$ bajo la hipótesis de igualdad de las cuatro varianzas. Esta hipótesis se puede poner a prueba mediante la prueba *Cochran*, por ejemplo, como en el numeral 2. Para una varianza máxima de 2,33 y una suma de todas las varianzas de 5,66, el estadístico de Cochran es $(2,33/5,66) = 0,41$. En comparación con el valor crítico 0,768 al 5 % para los cuatro laboratorios y tres réplicas, la igualdad de la varianza no se puede rechazar. En consecuencia, $s^2_r = 1,42$ con ocho grados de libertad.

Hay cuatro medias (16,00, 14,67, 14,33 y 15,00), una para cada laboratorio, en donde cada media es un estimado de la característica medida. El promedio total, m , es 15,00 y la varianza es 0,52 con tres grados de libertad. Debido a que en este caso cada media de laboratorio es un promedio de tres resultados sencillos, la varianza entre medias no es un estimado de s^2_L solamente, sino que contiene además parte de la varianza de repetibilidad (un tercio, en este ejemplo).

Esta información se puede resumir como en la Tabla 4:

Tabla 4. Estimación de los componentes de varianza

Cantidad estimada	Estimado
$s^2_L + (s^2_r)/3$	0,52
s^2_r	1,42

Igualando los estimados a las cantidades estimadas correspondientes, se obtienen s^2_r y s^2_L como estimados de s^2_r y s^2_L .

En donde:

s^2_r es igual a 1,42;

s^2_L es igual a 0,05 (por ej: $0,52 - 1,42/3$).

El estimado s_R^2 de la varianza de reproducibilidad está dado por $s_L^2 + s_r^2$ y es igual a 1,47 (por ej: 0,05 + 1,42).

La mayoría de paquetes de software reportará la siguiente tabla equivalente (Tabla 5):

Tabla 5. Estimación de los componentes de la varianza

Cantidad estimada	Estimado
$3s_L^2 + s_r^2$	1,56 (= 3 x 0,52)
s_r^2	1,42

Los cálculos se hacen más complicados cuando el número de resultados por laboratorio no es igual, y en este caso se recomienda el uso de un paquete de software. Sin embargo, el principio sigue siendo el mismo. Se presentará en los ejemplos desarrollados.

4.3.2 Ejemplo 2

La Tabla 6 ilustra los datos sin procesar para diferentes laboratorios y las medias y varianzas asociadas.

Tabla 6. Ejemplo 1. Medias y varianzas

Parámetro	Lab 1	Lab 2	Lab 3	Lab 4
Resultados	63	44	50	53
	57	51	40	57
	54	43	42	46
Media	58	46	44	52
$s_{r,i}^2$ ^a	21	19	28	31
Grados de libertad	2	2	2	2
^a El subíndice <i>i</i> indica el <i>i</i> -ésimo laboratorio				

De estos resultados, se deduce que:

- Media total, m = 50,00 [por ejemplo: (58 + 46 + 44 + 52)/4]
- Varianza de la media = 40,0
- Varianza de repetibilidad, s_r^2 = 24,75 [por ejemplo: (21 + 19 + 28 + 31)/4]

En consecuencia,

- Varianza entre laboratorios, s_L^2 = 31,75 (por ejemplo: 40 - 24,75/3)
- Varianza de reproducibilidad, $s_R^2 = s_L^2 + s_r^2$ = 56,50 (por ejemplo: 31,75 + 24,75)

4.4 USO DE LÍMITES DE REPETIBILIDAD Y REPRODUCIBILIDAD

La estimación de las desviaciones estándar de repetibilidad y reproducibilidad se pueden usar para obtener valores para los límites de repetibilidad y reproducibilidad. Estos límites se usan en la práctica para juzgar si es probable que el método de medición haya sido llevado a cabo sin efectos adversos, evaluar si dos productos se diferencian significativamente, o ambos.

Los límites de repetibilidad y reproducibilidad representan la diferencia absoluta entre dos únicos resultados obtenidos en condiciones de repetibilidad y reproducibilidad, respectivamente, que no excederán más del 95 % del tiempo. Una estimación aproximada de los límites de repetibilidad y reproducibilidad se puede obtener multiplicando las desviaciones estándar de repetibilidad y reproducibilidad, respectivamente, por 2,8 (ISO 5725-6:1994, numeral 4.1) y aplicando las siguientes reglas:

- a) Si dos resultados obtenidos del mismo producto en condiciones de repetibilidad (o condiciones de reproducibilidad) se diferencian en más del límite de repetibilidad (o reproducibilidad), es probable que ocurra un problema durante la aplicación del método, el muestreo, o ambos. Se recomienda realizar investigaciones adicionales. Posiblemente sean necesarios más resultados.
- b) Si dos resultados obtenidos de cada uno de dos productos diferentes en condiciones de repetibilidad (o de reproducibilidad) se diferencian en más del límite de repetibilidad (o reproducibilidad), es razonable cuestionar si estos dos productos son de calidad diferente.

NOTA Las desviaciones estándar de repetibilidad y reproducibilidad y los límites son estimados de precisión y por tanto están sujetos a errores de estimación. Están previstos para uso como guías para evaluar la validez o pertinencia de los resultados producidos por el método de medición y no como criterios numéricos rígidos para descartar o validar resultados. En particular, el valor del factor 2,8 usado para obtener los límites de repetibilidad y reproducibilidad se basa en la suposición de una distribución normal determinada completamente antes de que se obtenga cualquier resultado, lo cual rara vez ocurre en la práctica. Muchos profesionales han encontrado adecuado usar el valor de tres para esta constante, sin pérdida de información. El sentido común debería prevalecer en todos los casos.

Los resultados de la aplicación de este estimado de evaluación al ejemplo 2 son:

- Límite de repetibilidad: $r = 2,8 \times s_r = 13,93$;
- Límite de reproducibilidad: $R = 2,8 \times s_R = 21,05$

La diferencia entre los dos resultados obtenidos del mismo producto en condiciones de repetibilidad (o reproducibilidad) no excederá de 13,93 (ó 21,05) en 95 casos de 100.

5. EJEMPLOS DESARROLLADOS MEDIANTE EL USO DE SOFTWARE ESTADÍSTICO

5.1 GENERALIDADES

Este numeral ilustra el grado en que se pueden obtener automáticamente gráficas y estadísticas mediante el uso programas de software estadístico. El objetivo no es dar publicidad ni respaldar paquetes específicos, sino hacer énfasis en las principales ventajas de

los cálculos automatizados, a saber: exactitud, rapidez y la capacidad de usar procedimientos que podrían ser demasiado complicados usando solamente una calculadora de bolsillo.

Sin embargo, existen algunas limitaciones de uso de los paquetes de software. No todos los paquetes poseen la capacidad de calcular todos los parámetros estadísticos requeridos, por ejemplo, los valores estadísticos h y k de *Mandel* o los valores estadísticos de *Grubbs*.

Dos de los ejemplos de la NTC 3529-2 (ISO 5725-2), Anexo B, se presentan aquí: NTC 3529-2 (ISO 5725-2), B.1, ejemplo 1, “Determinación del contenido de azufre del carbón”, y la NTC 3529-2 (ISO 5725-2), B.3, ejemplo 3: “Titulación termométrica del aceite de creosota”. Cuando sea aplicable, se resaltarán las secciones específicas de este anexo.

5.2 DETERMINACIÓN DEL CONTENIDO DE AZUFRE EN EL CARBÓN

5.2.1 Datos originales

En general, los datos se pueden registrar de acuerdo con el formato de la Tabla B.1 de la NTC 3529-2 (ISO 5725-2). Una columna contiene el número del laboratorio, y en este caso, los datos se encuentran en cuatro columnas, una por nivel. Véase la Tabla 7.

Tabla 7. Presentación de los datos (representan el contenido de azufre del carbón^{a,b})

Laboratorio	Nivel			
	1	2	3	4
1	0,71	1,20	1,68	3,26
1	0,71	1,18	1,70	3,26
1	0,70	1,23	1,68	3,20
1	0,71	1,21	1,69	3,24
2	0,69	1,22	1,64	3,20
2	0,67	1,21	1,64	3,20
2	0,68	1,22	1,65	3,20
3	0,66	1,28	1,61	3,37
3	0,65	1,31	1,61	3,36
3	0,69	1,30	1,62	3,38
4	0,67	1,23	1,68	3,16
4	0,65	1,18	1,66	3,22
4	0,66	1,20	1,66	3,23
5	0,70	1,31	1,64	3,20
5	0,69	1,22	1,67	3,19
5	0,66	1,22	1,60	3,18
5	0,71	1,24	1,66	3,27
5	0,69	—	1,68	3,24
6	0,73	1,39	1,70	3,27
6	0,74	1,36	1,73	3,31
6	0,73	1,37	1,73	3,29
7	0,71	1,20	1,69	3,27
7	0,71	1,26	1,70	3,24
7	0,69	1,26	1,68	3,23
8	0,70	1,24	1,67	3,25
8	0,65	1,22	1,68	3,26
8	0,68	1,30	1,67	3,26

^a Datos originales de la NTC 3529-2 (ISO 5725-2), B.1.2.
^b Las unidades de concentración son el porcentaje de fracción de masa.

Como alternativa, los datos se pueden ingresar en tres columnas, una para laboratorio, otra para nivel, otra para datos, en este caso el contenido de azufre, expresado en porcentaje de fracción másica. Un ejemplo de esta disposición se presenta en la Tabla 8 para los laboratorios 1 y 2 solamente. El archivo entero continúa de forma similar para la inclusión de los laboratorios 3 al 8.

Tabla 8. Presentación alternativa de datos (representan el contenido de azufre del carbón^{a,b}) para los laboratorios 1 y 2

Laboratorio	Nivel	Azufre
1	1	0,71
1	1	0,71
1	1	0,70
1	1	0,71
1	2	1,20
1	2	1,18
1	2	1,23
1	2	1,21
1	3	1,68
1	3	1,70
1	3	1,68
1	3	1,69
1	4	3,26
1	4	3,26
1	4	3,20
1	4	3,24
2	1	0,69
2	1	0,67
2	1	0,68
2	2	1,22
2	2	1,21
2	2	1,22
2	3	1,64
2	3	1,64
2	3	1,65
2	4	3,20
2	4	3,20
2	4	3,20
^a datos originales de la NTC 3529-2 (ISO 5725-2) B.1.2.		
^b Las unidades de concentración son en porcentaje de fracción másica		

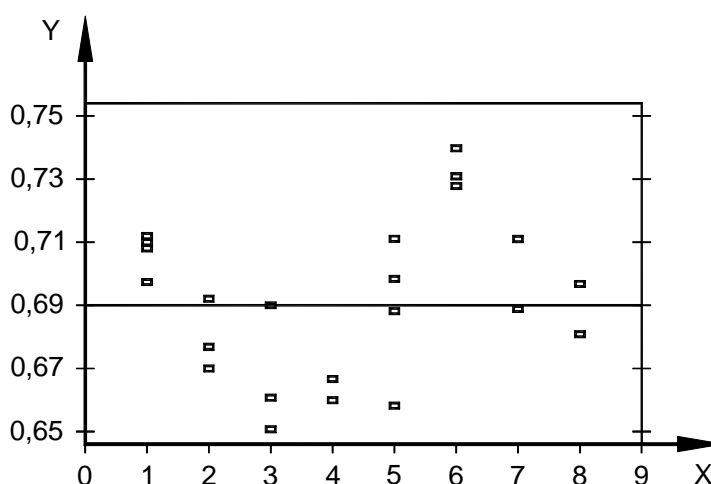
Como ya se indicó, se recomienda graficar los datos antes de realizar las pruebas, para detectar los valores atípicos potenciales. En seguida se presentan dos ejemplos de estos gráficos. La dispersión de los datos sin procesar para el nivel 1 se demuestra en la Figura 11, que es similar a la NTC 3529-2 (ISO 5725-2), Figuras B.1 a B.4. Estos gráficos muestran los resultados por laboratorio, con una línea horizontal trazada en el valor de la media total. El gráfico de “caja y bigotes” de la Figura 12, compuesto de una caja, “bigotes” y valores atípicos, resume los resultados por laboratorio para el nivel 1 y es útil para obtener una evaluación inicial de los valores atípicos potenciales. Es una visualización interesante e informativa de los datos.

El promedio y la mediana para cada laboratorio están representados respectivamente por una cruz y una línea más oscura. La parte inferior de la caja está en el primer cuartil (Q1), y la parte superior está en el tercer cuartil (Q3). El rango aproximado de datos se indica por las líneas de la parte superior y de la parte inferior de la caja, llamadas “bigotes”, que se extienden desde la parte superior e inferior de la caja a las observaciones más altas y más bajas, respectivamente, que siguen estando dentro de la región definida por los siguientes límites:

- Límite inferior: $Q1 - 1,5 (Q3 - Q1)$
- Límite superior: $Q3 + 1,5 (Q3 - Q1)$

Para una distribución normal o *gaussiana*, esto corresponde a aproximadamente al 99,30 % del área bajo la curva de probabilidad. *John Turkey*, inventor del gráfico de “caja y bigotes”, llama a estos límites las “cercas internas”. Definió además las “cercas externas” obtenidas usando un multiplicador de 3 en lugar de 1,5. Los puntos de datos que están dentro de las cercas “interna” y “externa” se definen como posibles valores atípicos. Los datos que se encuentran más allá de la cerca “externa” son bastante sospechosos, debido al hecho de que el área dentro de las cercas externas abarca el 99 % del área debajo de la curva. Algunos paquetes de software estadísticos usan diferentes símbolos para estas dos clases de valores atípicos, tales como “*” para los datos entre las cercas internas y externas y “0” para datos más allá de las cercas externas.

Ambos gráficos muestran que todos los resultados para el laboratorio 6 son más altos que los valores correspondientes para los otros laboratorios. Además, hay un valor para el valor 5 identificado fácilmente en la Figura 12, que podría ameritar investigación.



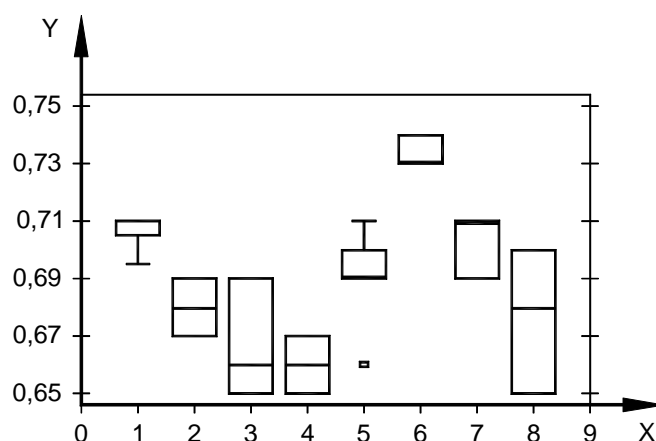
Convenciones

X Número del laboratorio

Y Concentración, fracción másica en porcentaje

- $m_1 = 0,69$

Figura 11. Gráfico de dispersión de datos nivel 1



Convenciones

X Número de laboratorio
Y concentración, fracción másica en porcentaje

? es la mediana
+ es el promedio
- es un "bigote"

Figura 12. Gráfico de "caja y bigotes" de datos nivel 1

5.2.2 Tabla de medias de celdas y desviaciones estándar

La Tabla 9 presenta resultados de la NTC 3529-2 (ISO 5725-2), B.1.3, y B.1.4, calculados para una prueba interlaboratorio con un paquete de software típico.

Tabla 9. Tabla de medias y desviaciones estándar de celdas

Número del laboratorio	Cálculo	Media	Desviación estándar
Total			
—	27	0,690 37	0,025 49
Para laboratorios individuales			
1	4	0,707 50	0,005 00
2	3	0,680 00	0,010 00
3	3	0,666 67	0,020 82
4	3	0,660 00	0,010 00
5	5	0,690 00	0,018 71
6	3	0,733 33	0,005 77
7	3	0,703 33	0,011 55
8	3	0,676 67	0,025 17

Como se ilustra en la NTC 3529-2 (ISO 5725-2), Tablas 2 y 3, estos valores se requieren para calcular los estadísticos h y k de *Mandel*, al igual que la prueba de *Grubbs*, que no se encuentran en todos los paquetes de software.

5.2.3 Escrutinio para la determinación de compatibilidad y de valores atípicos

En muchos paquetes hay varias pruebas disponibles para evaluar la homogeneidad de las varianzas, incluidas las de *Cochran*, *Bartlett*, *Hartley* y *Levene*. Los resultados de estas pruebas en datos para contenido de azufre se ilustran en la Tabla 10.

Tabla 10. Escrutinio para determinación de compatibilidad y valores atípicos

Prueba	Estadístico	Valor <i>P</i>
Cochran C	0,350	0,308
Bartlett	1,679	0,296
Levene	1,679	0,332
Hartley	25,333	—

El valor *P* representa la probabilidad de observar un estadístico tan grande como el que se calcularía si las varianzas fueran homogéneas. Un valor de *P* de 0,05 ó inferior es suficiente para concluir que las varianzas interlaboratorio son diferentes. En consecuencia, para la prueba de *Cochran* existe evidencia insuficiente para concluir que las varianzas intralaboratorio son diferentes.

5.2.4 Cálculo de la media general y de las desviaciones estándar de repetibilidad y reproducibilidad

Como se describe en la NTC 3529-2 (ISO 5725-2) y en la ISO 5725-3, el análisis de la técnica de varianza se puede usar para calcular s_{rj} y s_{Rj} , las desviaciones estándar de repetibilidad y reproducibilidad, respectivamente, para el nivel *j*.

La mayoría de paquetes tiene capacidad para calcular directamente los componentes de varianza. Esta capacidad se puede encontrar usualmente entre otras herramientas estadísticas buscando la opción titulada “Componentes de varianza” o en forma equivalente “Diseños jerárquicos” o “Diseños anidados”.

El anidado ocurre cuando las muestras analizadas son diferentes de un laboratorio a otro, aun cuando provengan de un solo lote o carga de producto. Por tanto, las muestras se anidan dentro del “laboratorio” o en forma equivalente, los laboratorios están jerárquicamente sobre las muestras.

Si no existe esta opción, otra opción es usar un análisis de varianza en un sentido, en donde el factor “laboratorio” (el factor en “un sentido”) es declarado aleatorio.

Los componentes de varianza hacen referencia a la partición de la varianza total en dos partes, las varianzas entre laboratorios y en un laboratorio, s_{Lj}^2 y s_{rj}^2 respectivamente. Un ejemplo del análisis de varianza resultante en un sentido, se da en la Tabla 11, para nivel 1 (*j* = 1).

Tabla 11. Análisis de componentes de varianza. Análisis de varianza del contenido de azufre. Nivel 1

Fuente	Suma de cuadrados	Grados de libertad	Cuadrado medio	Componente de varianza	Porcentaje
Laboratorio	0,012 554 6	7	0,001 793 5	0,000 466 5	67,1
Error	0,004 341 7	19	0,000 228 5	0,000 228 5	32,8
Total (corregido)	0,016 896 3	26	—	—	—

La desviación estándar de repetibilidad, s_{r1} , se obtiene tomando la raíz cuadrada de la componente de varianza “error”. En consecuencia:

$$s_{r1} = \sqrt{0,000\ 228\ 5} = 0,015$$

La desviación estándar de reproducibilidad, s_{R1} se obtiene tomando la raíz cuadrada de la suma del componente de varianza “error” y del componente de varianza “laboratorio”. En consecuencia:

$$s_{R1} = \sqrt{0,000\ 228\ 5 + 0,000\ 466\ 5} = 0,026$$

El peso relativo del componente de varianza “laboratorio” y del componente de varianza “error” para la varianza de reproducibilidad se indica en la columna “porcentaje”.

Si esta opción no estuviera disponible, entonces la tabla de análisis de varianza podría no visualizar directamente los componentes de la varianza y los resultados podrían aparecer como se ilustra en la Tabla 12. Estos valores se pueden obtener mediante cálculos manuales usando la fórmula dada en la NTC 3529-2 (ISO 5725-2), 7.4.5.2, como se ilustra en seguida:

Tabla 12. Tabla de análisis de varianza para contenido de azufre. Nivel 1. Análisis de varianza

Fuente	Suma de cuadrados	Grados de libertad	Cuadrados medios	Relación F	Valor P
Entre laboratorios	0,012 554 6	7	0,001 793 5	7,85	0,000 2
En laboratorio (error)	0,004 341 7	19	0,000 228 5	—	—
Total (corregido)	0,016 896 3	26	—	—	—

El parámetro s_{r1}^2 está dado por el valor del cuadrado medio en laboratorio: $s_{r1}^2 = 0,000\ 228\ 5$

Como se ilustra en la NTC 3529-2 (ISO 5725-2), numeral 7.4.5.2, ecuación (21)

$$S_{L1}^2 = \frac{s_{d1}^2 - s_{r1}^2}{n_1}$$

en donde

s_{d1}^2 es el cuadrado medio entre laboratorios igual a 0,001 793 52

$s_{r1}^2 = 0,000\ 228\ 5$ (véase arriba)

$$n_1 = \frac{1}{p_{Lab} - 1} \left(\sum n_{il} - \frac{\sum n_{il}^2}{\sum n_{il}} \right) \quad [\text{ISO 5725-2:1994, numeral 7.4.5.2, Ecuación (23)}]$$

$$= \frac{1}{7} \left(\frac{4^2 + 3^2 + 3^2 + 3^2 + 5^2 + 3^2 + 3^2 + 3^2}{27} \right) = 3,35$$

NOTA Una aproximación de \bar{n}_j se obtiene mediante el número medio de resultados por laboratorio (es decir, $27/8 = 3,375$).

En consecuencia:

$$s_{L1}^2 = \frac{0,0017935 - 0,0002285}{3,35} = 0,0004672$$

Ya que $s_{R1}^2 = s_{L1}^2 + s_{r1}^2$, entonces:

$$s_{R1}^2 = \sqrt{0,00022859 + 0,0004672} = 0,026$$

5.2.5 Dependencia de la precisión respecto a m

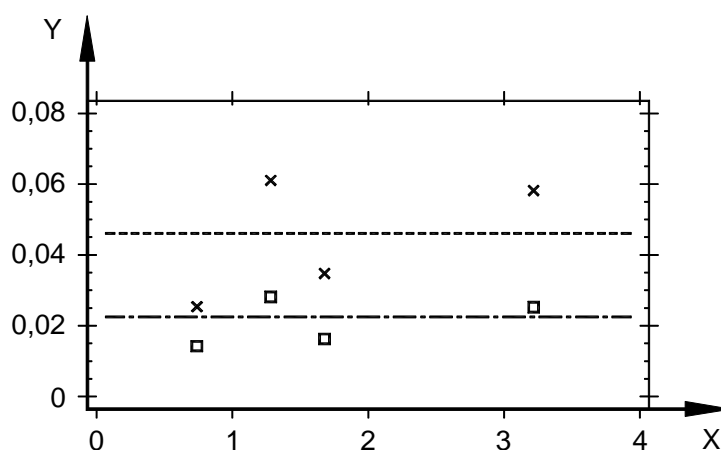
Cuando el análisis estadístico se repite para todos los otros niveles, se determinan desviaciones estándar de repetibilidad y reproducibilidad.

Los datos se pueden presentar como se indica en la Tabla 13

Tabla 13. Desviaciones estándar de reproducibilidad y repetibilidad para niveles múltiples

Nivel	Medias totales	Desviación estándar de repetibilidad	Desviaciones estándar de reproducibilidad
1	0,690	0,015	0,026
2	1,252	0,029	0,061
3	1,667	0,017	0,035
4	3,250	0,026	0,058

La dependencia de la precisión con respecto a la media total del nivel (véase también la NTC 3529-2 (ISO 5725-2), B.1.7) se puede evaluar observando la gráfica de dispersión que se presenta abajo. Las desviaciones estándar de repetibilidad y reproducibilidad promedio, 0,022 y 0,045 respectivamente, se calculan y se pueden trazar como se ilustra en la Figura 13.



Convención

X media total en porcentaje de fracción másica
Y porcentaje de fracción másica

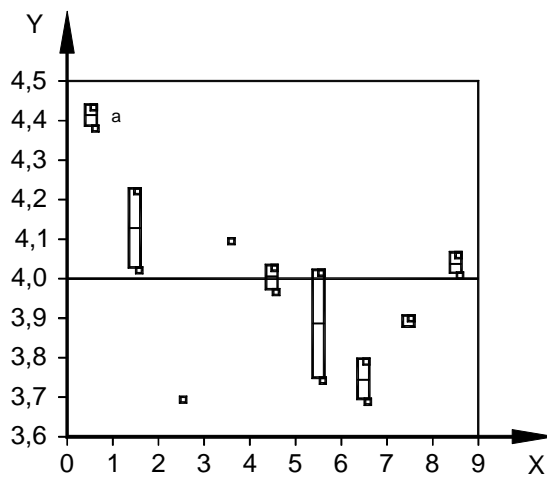
repetibilidad
X reproducibilidad
... promedio s_R
-.-.- promedio s_x

Figura 13. Gráfico de s_x y s_R contra la media total

5.3 TITULACIÓN TERMOMÉTRICA DEL ACEITE DE CREOSOTA

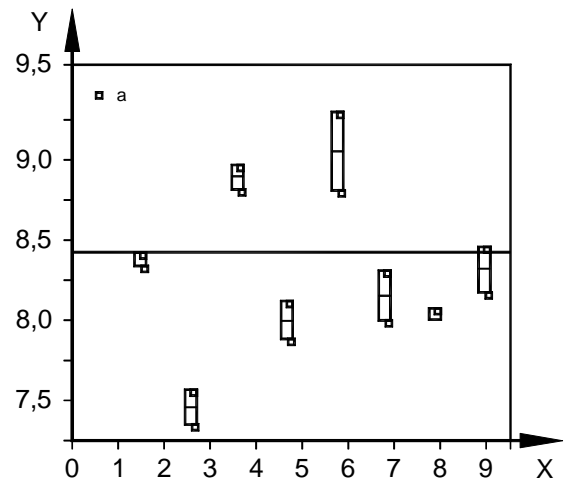
5.3.1 Datos originales

Los gráficos de los datos sin procesar de la NTC 3529-2 (ISO 5725-2):1994, B.3.2 para la titulación termométrica del aceite de creosota muestran que el laboratorio 1 siempre produce datos que son constantemente más altos que los de todos los otros laboratorios. Este hecho podría ameritar investigación adicional. Además, las pruebas de valores atípicos podrían sugerir que estos datos se deberían considerar para rechazo.



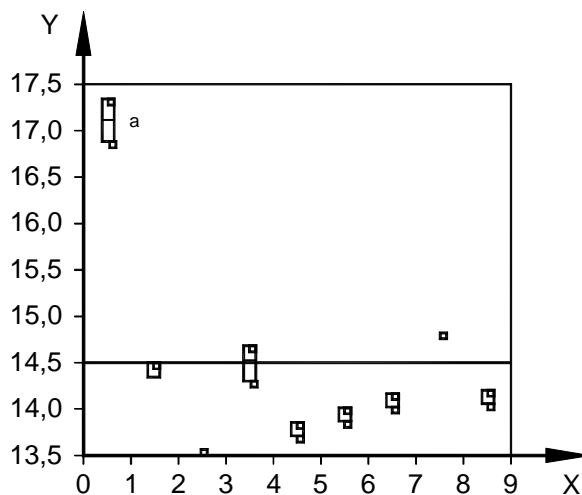
a) Nivel 1 de chequeo de varianza

Prueba C de Cochran: 0,566 474 P-valor = 0,218 148
 Prueba de Bartlett: 3,839 36 P-valor = 0,337 881
 Prueba de Hartley: 784,0



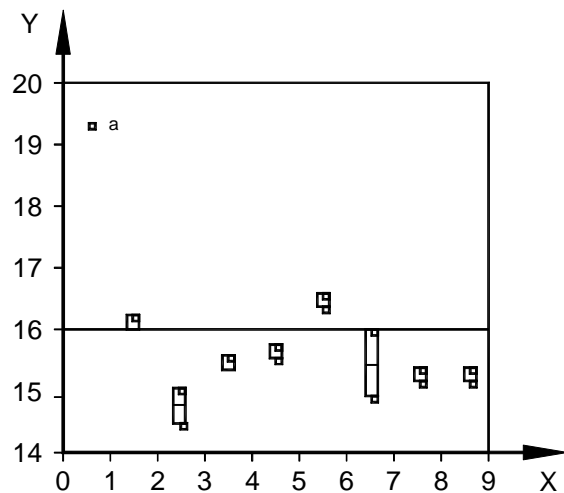
b) Nivel 2 de chequeo de varianza

Prueba C de Cochran: 0,449 912 P - valor = 0,383 8
 Prueba de Bartlett: 2,274 14 P - valor = 0,686 763
 Prueba de Hartley: 256,0



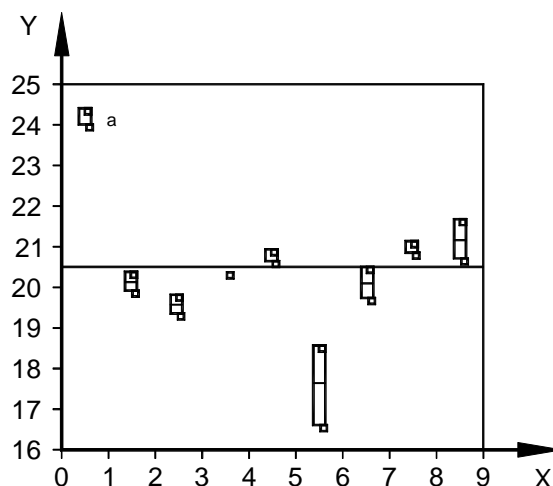
c) Nivel 3 de chequeo de varianza

Prueba C de Cochran: 0,492 417 P – valor = 0,366 708
 Prueba de Bartlett: 2,171 5 P – valor = 0,686 072
 Prueba de Hartley: 39,062 5



d) Nivel 4 de chequeo de varianza

Prueba de C Cochran: 0,666 703 P-valor = 0,057 952 4
 Prueba de Bartlett: 3,271 86 P – valor = 0,439 721
 Prueba de Hartley: 121,0



e) Nivel 5 de chequeo de varianza

Prueba C de *Cochran*: 0,635 778

P – valor = 0,080 422 1

Prueba de *Bartlett*: 2,485 46

P – valor = 0,623 743

Prueba de *Hartley*: 50,005 1

Convenciones

X	Número de laboratorio
Y	creosota, fracción másica en porcentaje
–	promedio de todos los promedios de laboratorio
a	Valor sistemáticamente alto

Figura 14. Gráficos de “caja y bigotes” y pruebas de varianza

Los gráficos de “caja y bigotes” ilustran claramente que el laboratorio 1 presenta un valor sistemático más alto que los otros laboratorios, aunque los laboratorios 6 y 7 muestran una mayor variabilidad en algunos niveles específicos.

5.3.2 Escrutinio para determinación de compatibilidad y valores atípicos

Los datos se pueden presentar en el formato de la Tabla 14 para evaluar la compatibilidad y la identificación de los valores atípicos; véase también la NTC 3529-2 (ISO 5725-2), B.3.2 y B.3.5.

Tabla 14. Escrutinio para determinación de compatibilidad y valores atípicos

Laboratorio	Observación	Media por nivel				
		Nivel 1	Nivel 2	Nivel 3	Nivel 4	Nivel 5
1	alto	4,415	9,340	17,150	19,230	24,140
2	–	4,130	8,375	14,460	16,140	20,155
3	–	3,700	7,500	13,600	14,800	19,500
4	–	4,100	8,865	14,400	15,550	20,300
5	–	4,005	8,005	13,825	15,660	20,700
6	–	3,890	9,000	13,980	16,500	17,570
7	–	3,750	8,150	14,150	15,450	20,100
8	–	3,905	8,055	14,840	15,315	20,940
9	–	4,045	8,305	14,170	15,290	21,185
Parámetro		Valor por nivel				
Media		3,993	8,399	14,508	15,993	20,511
Desviación estándar		0,216	0,572	1,056	1,310	1,727
Media del error estándar		0,072	0,191	0,352	0,437	0,576

EJEMPLO

para el nivel 3 y el laboratorio 1, el único estadístico de *Grubbs* alto, G_b , está dado por:

$$G_b = \frac{17,15 - 14,508}{1,056} = 2,50$$

Los valores críticos en el nivel de significación del 5 % y del 1 % son 2,215 y 2,387 respectivamente, para los nueve laboratorios. En consecuencia, con una confianza del 99 %, los datos para nivel 3 y laboratorio 1 se pueden considerar valores atípicos.

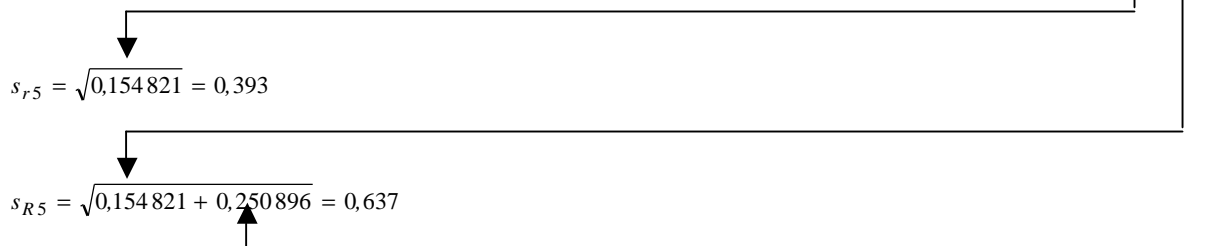
5.3.3 Cálculos de la media total y de las desviaciones de repetibilidad y reproducibilidad

Con esta base, todos los datos del laboratorio 1, junto con los datos del laboratorio 6 a nivel 5 se rechazan. De acuerdo con esto, el análisis de varianza para el nivel 5 ($j = 5$) y los cálculos de las desviaciones estándar de repetibilidad y reproducibilidad, s_r y s_R respectivamente, se dan como se ilustra en la Tabla 15 y en las ecuaciones siguientes; véase también la NTC 3529-2 (ISO 5725-2), B.3.6.

Tabla 15. Análisis de varianza de nivel 5 para datos de la titulación termométrica de la creosota

Análisis de varianza						Estimados de componentes de varianzas ^a	
Fuente	Grados de libertad	Suma de cuadrados	Cuadrado medio	Relación F	Probabilidad (F > 4,24)	Componente	Estimado
Modelo	6	3,939 686	0,656 614	4,241 1	0,040 1	Laboratorio	0,250 896
Residual	7	1,083 750	0,154 821	–	–	Residual	0,154 821
Total	13	5,023 436	–	–	–		

^a Estos estimados se basan en igualar cuadrados medios al valor esperado



El cuadrado medio del modelo es un estimado de $2 \sigma_L^2 + \sigma_r^2$

El cuadrado medio residual es un estimado de σ_r^2

5.3.4 Dependencia de la precisión respecto a m

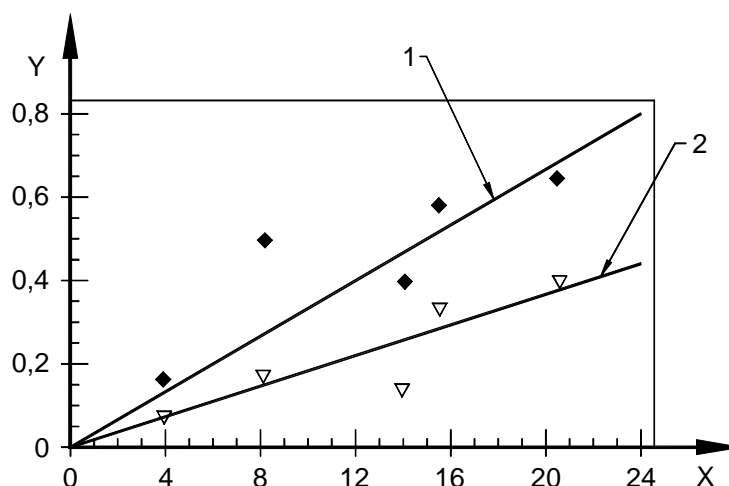
Después de rechazar los valores atípicos, se recalculan nuevas medias totales, y a partir de ellas, se calculan las desviaciones estándar de repetibilidad y reproducibilidad como se indica en la Tabla 16; véase también la NTC 3529-2 (ISO 5725-2), B.3.7.

Tabla 16. La precisión en función de la media total para datos de la titulación termométrica de la creosota

Nivel	Media total	Desviación estándar de repetibilidad	Desviación estándar de reproducibilidad
1	3,94	0,092	0,171
2	8,28	0,179	0,498
3	14,18	0,127	0,400
4	15,59	0,337	0,579
5	20,41	0,393	0,637

La dependencia de la precisión respecto a las medias totales del nivel se puede evaluar observando el diagrama de dispersión de la Figura 15.

En este caso, un análisis de regresión lineal es más apropiado (la pendiente es significativamente diferente de 0) que las desviaciones estándar de repetibilidad y reproducibilidad promediadas, y la relación es como se indica en la Figura 15:



Convenciones

X Media total, m , expresada como una fracción de masa

Y Desviación estándar, expresada como una fracción de masa

1 Desviación estándar de reproducibilidad $s_R = 0,034 m$

2 Desviación estándar de repetibilidad $s_r = 0,018$

Figura 15. Curva de s_r y s_R para datos de la titulación termométrica de la creosota

Las Tablas 17 y 18 presentan resultados de la regresión para las desviaciones estándar de repetibilidad y reproducibilidad, respectivamente. Para propósitos de ilustración, aunque esta relación no tiene sentido desde un punto de vista físico-químico, se ha adaptado una línea recta a través del origen, como se ilustra en la Figura 15. El ajuste de una línea recta es complicado debido al hecho de que los valores de la media y de la precisión son estimados y por tanto están sujetos a error. Un buen estimado de los parámetros de la línea de regresión requiere una regresión ponderada que no se ha realizado debido a que la mayoría de tiempo la pendiente de la línea recta es pequeña.

Tabla 17. Desviación estándar de repetibilidad para datos de la titulación termométrica de la creosota

Análisis de regresión ^a					
Parámetro	Estimado	Error estándar	Estadística t	Probabilidad $t > 7,49$	
Media por nivel	0,017 909 6	0,002 391 7	7,488 62	0,001 7	
Fuente	Suma de cuadrados	Grados de libertad	Cuadrado medio	Relación F	Probabilidad F > 73,94
Modelo	0,303 037	1	0,303 037	56,08	0,001 7
Residual	0,021 615	4	0,005 404	—	—
Total	0,324 652	5	—	—	—
Raíz media cuadrática = 0,073 510					
Media absoluta residual = 0,052 872					
^a La variable dependiente es la desviación estándar de repetibilidad					

Tabla 18. Desviación estándar de reproducibilidad para datos de la titulación termométrica de la creosota

Análisis de regresión ^a					
Parámetro	Estimado	Error estándar	Estadística <i>t</i>	Probabilidad <i>t</i> > 7,49	
Media por nivel	0,034 393 7	0,004 000 1	8,599	0,001 0	
Fuente	Suma de cuadrados	Grados de libertad	Cuadrado medio	Relación F	Probabilidad F > 73,94
Modelo	1,117 790	1	1,117 790	73,94	0,0010
Residual	0,060 468	4	0,015 117	—	—
Total	1,178 258	5	—	—	—
Raíz media cuadrática = 0,122 951					
Media absoluta residual = 0,088 842					
^a La variable dependiente es la desviación estándar de reproducibilidad					

Ya que los valores *P* son inferiores a 0,05, la dependencia que tiene la precisión respecto a la media total para un nivel particular se puede expresar razonablemente por:

$$s_r = 0,018 \, m$$

$$s_R = 0,034 \, m$$

Estas relaciones se pueden usar para estimar las desviaciones estándar de repetibilidad y reproducibilidad a niveles intermedios dentro del rango de niveles estudiados.

Por ejemplo, si las desviaciones estándar de repetibilidad y reproducibilidad se requieren para una media total, *m*, en donde *m* = 12, entonces:

$$S_r = 0,018 \times 12 = 0,22$$

$$S_R = 0,034 \times 12 = 0,41$$

ANEXO A
(Normativo)

SÍMBOLOS Y ABREVIATURAS

Todos los símbolos y abreviaturas usados en este documento son los mismos de la ISO 5725 (todas las partes), con excepción de lo siguiente:

p_{Lab} ha reemplazado a p para diferenciarlo claramente de P , el símbolo para “probabilidad”

a	Intercepto en la relación $s = a + bm$
A	Factor usado para calcular la incertidumbre de un estimado
b	Pendiente en la relación $s = a + bm$
B	Componente en el resultado de una prueba, que representa la desviación de un laboratorio en relación con el promedio general (componente de sesgo de laboratorio)
B_0	Componente de B que representa todos los factores que no cambian en condiciones de precisión intermedia
$B(1), B(2)$, etc.	Componente de B que representa factores que varían en condiciones de precisión intermedia
c	Intercepto en la relación $\lg s = c + d \lg m$
C, C', C''	Estadístico de prueba
$C_{crit}, C_{crit}', C_{crit}''$	valores críticos para pruebas estadísticas
CD_P	diferencia crítica para la probabilidad P
CR_P	rango crítico para probabilidad P
d	pendiente en la relación $\lg s = c + d \lg m$
e	componente en un resultado de prueba, que representa el error aleatorio que ocurre en cada resultado de prueba
f	factor de rango crítico
$F_p(v_1, v_2)$	cuantilo p de la distribución F con v_1 y v_2 grados de libertad
G	estadístico de la prueba de <i>Grubbs</i>

h	estadístico de una prueba de <i>Mandel</i> de compatibilidad entre laboratorios
k	estadístico de una prueba de <i>Mandel</i> de compatibilidad en laboratorio
LCL	nivel de control inferior (ya sea la acción límite o el límite de advertencia)
m	media general de la propiedad de prueba; nivel
M	número de factores considerados en condiciones de precisión intermedia
n	numero de resultados de las pruebas obtenidos por laboratorio al nivel uno (es decir, por celda)
N	número de repeticiones
p_{Lab}	número de laboratorios que participan en un experimento entre laboratorios
	NOTA El símbolo usado en la serie ISO 5725 es p
P	probabilidad
q	número de niveles de la propiedad de la prueba en un experimento entre laboratorios
r	límite de repetibilidad
R	límite de reproducibilidad
MR	Material de referencia
s	estimado de una desviación estándar
\hat{s}	desviación estándar prevista
t	número de objetos o grupos de prueba
T	total o suma de alguna expresión
LCS	límite de control superior (ya sea límite de acción o límite de advertencia)
w	rango de un conjunto de resultados de prueba
W	factor de ponderación usado para calcular una regresión ponderada
x	referencia usada para la prueba de <i>Grubbs</i>
y	resultado de la prueba
\bar{y}	media aritmética de los resultados de prueba
\bar{y}	media <i>Frans</i> de los resultados de prueba

α	nivel de significación
β	probabilidad de error tipo II
γ	relación de la desviación estándar de reproducibilidad con la desviación estándar de repetibilidad (σ_R/σ_r)
Δ	sesgo de laboratorio
$\hat{\Delta}$	estimado de Δ
d	sesgo del método de medición
\hat{d}	estimado de δ
λ	diferencia detectable entre dos sesgos de laboratorio o los sesgos de dos métodos de medición
μ	valor real del valor de referencia aceptado de una propiedad de una prueba
ν	número de grados de libertad
s	valor real de una desviación estándar
t	componente en un resultado de una prueba, que representa la variación debida al tiempo desde la última calibración
f	relación detectable entre las raíces cuadradas de los cuadrados medios entre laboratorios del Método B y A
$x^2_p(\nu)$	cuantilo p de la distribución x^2 con ν grados de libertad

Símbolos usados como subíndices

C	calibración diferente
E	equipo diferente
i	identificador de un laboratorio particular
$I()$	identificador para medidas intermedias de precisión; entre paréntesis, identificación del tipo de situación intermedia
j	identificador para un nivel particular (NTC 3529-2 (ISO 5725-2)) identificador para un grupo de pruebas o para un factor (ISO 5725-3)
k	identificador para un resultado de una prueba particular en un laboratorio i a nivel j
L	entre laboratorios (interlaboratorio)
m	identificador de sesgo detectable

<i>M</i>	muestra entre pruebas
<i>O</i>	operador diferente
<i>P</i>	probabilidad
<i>r</i>	repetibilidad
<i>R</i>	reproducibilidad
<i>T</i>	hora diferente
<i>W</i>	intralaboratorio
1,2,3,....	para resultados de pruebas, numeración en el orden en el que se obtuvieron
(1), (2), (3)	para resultados de pruebas, numeración en el orden de magnitud creciente

BIBLIOGRAFÍA

- [1] ISO 5725-1:1994, Accuracy (trueness and precision) of measurement methods and results – Part 1: General principles and definitions (NTC 3529-1).
- [2] ISO 5725-2: 1994, Accuracy (trueness and precision) of measurement methods and results – Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method (NTC 3529-2).
- [3] ISO 5725-3: 1994, Accuracy (trueness and precision) of measurement methods – Part 3: Intermediate measures of the precision of a standard measurement method (NTC 3529-3).
- [4] P. DAGNÉLIE, 1975, Théorie et méthodes statistiques, Les Presses Agronomiques de Gembloux
- [5] SHAYLE R. SEARLE, George CASELLA, Charles E. McCULLOCH, 1992, variance Components, Wiley-Interscience.
- [6] MILLIKEN and JOHNSON, Analysis of messy data, Vol I Design Experiments, Van Nostrand Reinhold Company.
- [7] George E.P. Box, William G. HUNTER, J. STUART Hunter, Statistics for Experimenters. An introduction to design, data analysis and model building. John Wiley and Sons.
- [8] Vic BARNETT and Toby LEWIS, Outliers in statistical data, John Wiley and Sons

Programas de software estadístico¹

- | | | |
|------|--------------|--|
| S[1] | SAS | SAS Institute, Inc.
SAS Campus Drive, Cary, NBC 27513 |
| S[2] | Statgraphics | Manugistics Inc.
2215 East Jefferson St. Rockville, MD 20582 |
| S[3] | Statistica | Statsoft
2300 East 14 th St, Tulsa, OK 74104 |
| S[4] | Minitab | Minitab Inc.
3081 Enterprise Drive, State College, PA 16801-3008 |
| S[5] | SPSS | SPSS Inc.
44N Michigan Avenue, Chicago, IL 60611 |
| S[6] | S-PLUS | Insightful Corp.
1700 Westlake Avenue North, Suite 500, Seattle, WA 98109 |

¹ Esta bibliografía no es completamente extensa ni constituye una recomendación de los productos listados.

DOCUMENTO DE REFERENCIA

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. Accuracy (Trueness And Precision) of Measurement Methods and Results. Practical Guidance for the use of ISO 5725-2:1994 in Designing, Implementing and Statistically Analyzing Interlaboratory Repeatability and Reproducibility Results. Genève: ISO, 2005, 33 p. (ISO/TR 22971).