



**INTI**

**50**  
ANIVERSARIO  
1957 - 2007

Instituto  
Nacional  
de Tecnología  
Industrial



Unión Europea

Proyecto Mejora de la Eficiencia y de la Competitividad de la Economía Argentina

SERVICIO  
GEOLÓGICO  
MINERO  
ARGENTINO  
(SEGEMAR)

# CÁLCULO DE LA INCERTIDUMBRE EN LOS RESULTADOS ANALÍTICOS

CUADERNO TECNOLÓGICO N° 1  
**SEGEMAR**

Autor  
DR. F. XAVIER RIUS FERRUS

Junio de 2007

SERVICIO GEOLÓGICO MINERO  
ARGENTINO (SEGEMAR)



Instituto  
Nacional  
de Tecnología  
Industrial



Unión Europea

INTI

50  
ANIVERSARIO  
1957-2007

Proyecto Mejora de la Eficiencia y de la Competitividad de la Economía Argentina



Unión Europea

Delegación de la Comisión Europea en Argentina  
Ayacucho 1537  
Ciudad de Buenos Aires  
Teléfono (54-11) 4805-3759  
Fax (54-11) 4801-1594



INTI

Instituto  
Nacional  
de Tecnología  
Industrial

INTI  
Avenida General Paz 5445  
Casilla de Correo 157, B1650WAB San Martín, Buenos Aires, Argentina  
Teléfono (54-11) 4754-4070 / grobles@inti.gov.ar

[www.ue-inti.gov.ar](http://www.ue-inti.gov.ar)

#### CONTACTO

##### SERVICIO GEOLÓGICO MINERO ARGENTINO (SEGEMAR)

infointemin@inti.gov.ar  
www.segemar.gov.ar  
www.mineria.gov.ar  
Atención al cliente: 4754-4070  
Ing. Ramiro Fernández

INFORMACIÓN Y VISIBILIDAD: GUILLERMINA ROBLES  
grobles@inti.gov.ar

## CÁLCULO DE LA INCERTIDUMBRE EN LOS RESULTADOS ANALÍTICOS

DR. F. XAVIER RIUS FERRUS

# INDICE

## 0. PREÁMBULO

## 1. VALIDACIÓN DE MÉTODOS ANALÍTICOS

## 2. EXACTITUD Y TRAZABILIDAD

## 3. VERIFICACIÓN DE LA TRAZABILIDAD

- 3.1 Comparación de una media de un conjunto de resultados con un valor de referencia
- 3.2 Comparación de las medias de dos conjuntos de resultados. Conjuntos de resultados independientes entre sí
- 3.3 Comparación de dos conjuntos de resultados. Conjuntos de resultados relacionados o pareados
- 3.4 Comparación de dos conjuntos de resultados. Aplicación de la regresión lineal
- 3.5 Comparación de las medias de diversos conjuntos de resultados. Análisis de la varianza (ANOVA)

## 4. PRECISIÓN E INCERTIDUMBRE

## 5. CÁLCULO DE LA INCERTIDUMBRE

- 5.1 Cálculo de la incertidumbre mediante el método de la ISO
  - 5.1.1 Especificación
  - 5.1.2 Identificación
  - 5.1.3 Cuantificación
  - 5.1.4 Combinación
- 5.2 Cálculo de la incertidumbre basado en el proceso de verificación de la trazabilidad
  - 5.2.1 Incertidumbre asociada al procedimiento
  - 5.2.2 Incertidumbre asociada a la referencia
  - 5.2.3 Incertidumbre asociada a la muestra
  - 5.2.4 Cálculo de la incertidumbre del resultado final

## 6. APÉNDICE

- 6.1 Pruebas de significación

## 7. BIBLIOGRAFÍA

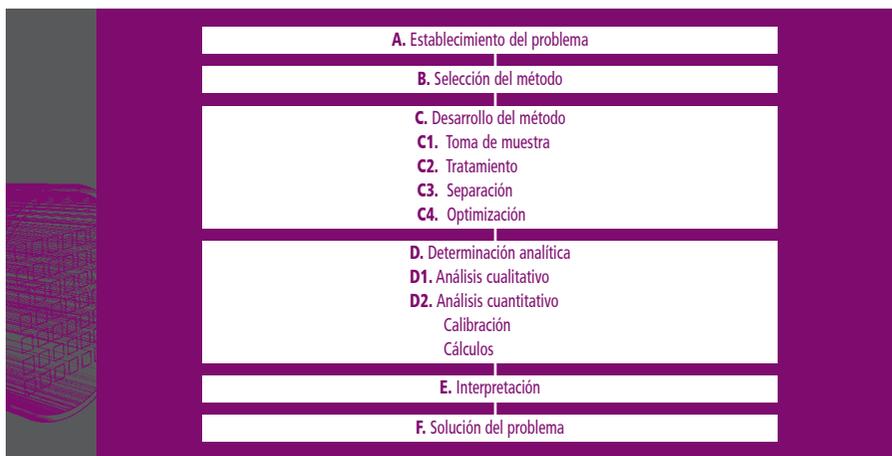
## EJERCICIOS PRÁCTICOS

## 0. INTRODUCCIÓN

El siguiente documento ha sido elaborado como complemento escrito al curso de especialización dictado en SEGEMAR-INTEMIN, entre el 4 al 15 de septiembre de 2006, bajo el título: Capacitación de profesionales y/o técnicos de laboratorios y Pymes y sus asociaciones en el uso de la normativa y gestión de la calidad. Tema: Conocimiento de las problemáticas en los ensayos de tipificación de minerales industriales y rocas de aplicación; y ensayos medioambientales, en cuanto a la calidad de los resultados de los laboratorios. Esta capacitación fue realizada por el Dr. F. Xavier Rius Ferrus, durante el mes de septiembre de 2006. El Dr. Rius Ferrus es Dr. en Química de la Universidad de Barcelona del año 1982. Es Profesor Catedrático de la Facultad de Química de la Universitat Rovira i Virgili desde el año 1991. Sus líneas de investigación son: Quimiometría, Cualimetría y Nanosensores Analíticos. Ha participado en numerosos Proyectos de I+D financiados en Convocatorias Públicas. Es autor de numerosas publicaciones en revistas internacionales sobre sus líneas de investigación. Ha participado en numerosos Proyectos de especial relevancia en Empresas y/o Administraciones. Es autor de numerosas contribuciones a Congresos Internacionales de la especialidad. Desde Mayo de 2002 es Director de la escuela de Postgrado y Doctorado de la Universitat Rovira i Virgili. Es miembro del Consejo editorial de Journal of Chemometrics y miembro del Consejo editorial de Analytica Chimica Acta.

Los contenidos de este documento completan los materiales (contenidos, gráficos, esquemas y tablas) proyectados durante las sesiones del curso. Dichas proyecciones, al estar incluidas como anexo al informe del curso, no se han incluido de nuevo en el presente documento.

Un resultado analítico se caracteriza por ser el producto de un proceso global de medida formado, normalmente, por diversas etapas independientes esquematizadas en el siguiente diagrama:



Esquema simplificado del proceso de medida químico

Este desarrollo particular de cualquier proceso de medida químico lleva consigo una serie de características que lo diferencian de otros procesos de medida, en los que intervienen exclusivamente magnitudes físicas.

Debido principalmente a que la variable final del análisis resulta ser la suma de un número relativamente grande de variables aleatorias, se puede aplicar al proceso el teorema del límite central y admitir como hipótesis válida que la distribución de los resultados es muy aproximadamente normal o Gaussiana. (HAHN, 1991), (BOX, 1978) Esta hipótesis es con la que trabajan casi la totalidad de los analistas más prestigiosos.

Por otra parte, en las diversas etapas esquematizadas en la figura, intervienen operaciones como la filtración, la precipitación, la evaporación, etc., en las que no es posible garantizar la conservación de las masas totales en juego y, como consecuencia, el concepto de trazabilidad parece desdibujarse casi totalmente para la magnitud de interés, que es la concentración de un analito en una matriz determinada.

El concepto de trazabilidad es uno de los más importantes para el analista y viene a complementar el concepto de exactitud (CITAC/EURACHEM 2000). Proveniente del campo de la metrología, se define formalmente como "la propiedad del resultado de una medición que consiste en que se pueda establecer el resultado previsible de su comparación directa con los patrones apropiados, por lo general, nacionales o internacionales, por medio de una cadena ininterrumpida de comparaciones reales". En otras palabras, el concepto de trazabilidad se relaciona con la propiedad de establecer las referencias adecuadas de nuestro resultado. Un resultado por sí solo no tiene mucho valor si no podemos acreditar su conexión (i.e., establecer la traza o trazar) con una referencia acreditada.

Debido a la distinta concepción de la trazabilidad en las mediciones físicas y las de naturaleza química, y para mantenernos dentro del espíritu de las normas europeas, debemos demostrar que es posible hacer análisis químicos trazables y con ese propósito tenemos que hacer una pequeña excursión mental. Pero, antes, conviene dejar bien establecida una propiedad que tenemos que exigir a nuestras medidas: su precisión, en lenguaje tradicional, su incertidumbre, en lenguaje más formal (aunque como se verá posteriormente, la incertidumbre engloba más aspectos del análisis químico y es más general que la precisión).

Hay que admitir que, si bien la conservación de masas no puede garantizarse en las operaciones que se han enumerado más arriba, las pérdidas o adiciones pueden hacerse repetitivas si se cuida la forma en que aquéllas se realizan. Se acepta que la demostración de que se ha cumplido esa condición se obtiene examinando la repetibilidad de los resultados finales de medida. Esta cualidad se ha descrito como precisión de las medidas y su valor puede ser cuantificado mediante diversas variables aleatorias (llamadas también estadísticos), construidas a partir de los datos experimentales. En este punto también es importante introducir el concepto de que al calcular la precisión se tiene que especificar las condiciones en las que ésta se ha calculado. Lógicamente no se obtendrán los mismos valores al calcular la precisión en condiciones de repetibilidad (diversas réplicas una detrás de otra utilizando el mismo instrumento, el mismo analista y el mismo laboratorio) que al calcular la precisión en condiciones de reproducibilidad (diversas réplicas en diversos días cambiando instrumento, analista e incluso laboratorio). Las condiciones de

repetibilidad dan lugar a los mínimos valores de precisión y las condiciones de reproducibilidad dan lugar a los máximos valores de precisión. Entre estas dos podemos encontrar las condiciones intermedias de tiempo (diversas réplicas en diversos días cambiando instrumento y analista pero dentro de un mismo laboratorio), que son las condiciones que tradicionalmente se conocen como reproducibilidad.

Nosotros hemos adoptado la varianza como el mejor indicador de la dispersión: a menor varianza, mayor precisión. La varianza experimental se obtiene de la manera (o sea, mediante la fórmula) presentada en la primera ponencia.

Interesa mucho destacar un detalle en el que posiblemente no se hace suficiente hincapié: las condiciones de repetibilidad de una medida se describen habitualmente mediante una lista de magnitudes de influencia que se conservan constantes. Magnitud de influencia, según el Vocabulario Internacional de Metrología, es aquella que no es objeto de la medición pero influye sobre el valor de medido o sobre las indicaciones del aparato a medir. Ejemplos de ella serían la temperatura ambiente o la humedad. Valga esa descripción como una primera aproximación descriptiva y gráfica pero, en rigor, esa constancia no es absoluta. Si lo fuera, no habría variabilidad y el resultado sería también invariable.

Las que consideramos como condiciones de repetibilidad en cualquier problema están definidas en realidad por la elección de aquellas magnitudes de influencia para las cuales decidimos no llevar nuestra indagación más allá y nos limitamos a tomar nota de su existencia a través de la varianza. Por así decir, la definición de las condiciones de repetibilidad depende de la forma en que se obtiene la varianza.

Obtenida pues, por definición, en condiciones de repetibilidad, este estadístico ofrece una imagen parcial de la calidad de las medidas. Hay que averiguar si esta imagen se repite para condiciones diferentes, es decir, si en condiciones de intermedias de tiempo, la dispersión de los resultados de las diferentes series de medidas es una característica estable del método de medida.

Si se puede demostrar esto último, mediante alguno de los tests estadísticos a nuestra disposición, también se puede comprobar si las medias de las diferentes series se parecen entre sí en el orden de dispersión que les correspondería, según ese valor de la varianza común. Cuando estas dos condiciones se cumplen, se dice que el sistema de medida; instrumentación, método y operador incluidos; está bajo control estadístico.

Así pues, el primer cuidado del analista químico es cerciorarse de que sus resultados se encuentran bajo control estadístico. Por todo lo dicho anteriormente, esta condición equivale a decir que cuando se aplica el sistema a la medida de un mensurando concreto, el resultado que se obtiene tiene un valor bien definido, al que se asocia una incertidumbre estadística, (o de tipo A), que puede obtenerse a partir de la desviación típica propia del laboratorio, en acuerdo completo con la forma de obtener la incertidumbre que se explicó en la ponencia anterior.

Una vez conseguido para el laboratorio ese estado, para completar la calidad de la medida es menester conocer el sesgo del resultado que se obtenga respecto del valor que para el ambiente en que nos movamos se considere convencionalmente como "verdadero". Si este sesgo fuese conocido, bastaría "corregir" por él el resultado y la medida ten-

dría, además de una precisión bien definida, una "exactitud" garantizada. Esta al menos sería la forma de expresión tradicional, que no es desdeñable, pero puede equivocar.

En lenguaje más actual, se diría que un resultado cuya corrección es conocida es trazable al valor convencionalmente verdadero (más rigurosamente, al valor de referencia).

Además, la disminución del énfasis que se obtiene hablando de trazabilidad a un valor convencionalmente verdadero determinado tiene la ventaja de sugerir por sí mismo el camino para una trazabilidad universal, aunque ese concepto sea por ahora utópico.

En efecto, en estos momentos, tratándose de análisis químicos, un valor convencionalmente verdadero es el resultado de la comparación de los resultados de diversos métodos, aplicados por diferentes laboratorios, en condiciones ambientales diversas, etc. En suma, es lo que resulta de aplicar a mayor escala el sistema utilizado para definir la incertidumbre típica de un laboratorio.

El salto mental es bien simple: imaginemos una comparación universal, con condiciones ambientales y de cualquier otra naturaleza variadas hasta el infinito y el resultado sería un valor convencionalmente verdadero con el mismo carácter que un patrón primario de una magnitud física fundamental. Mientras llega algo así, nos basta con manejarnos con valores convencionalmente verdaderos válidos para el ambiente en que nos movamos.

Definida la trazabilidad de los análisis químicos, asociándola a una versión mejorada de la exactitud de los resultados, falta exponer el modo en que, en la práctica, se consigue la trazabilidad.

Esa realización práctica no es sino el modo de llevar a cabo la comparación de resultados entre laboratorios. Por supuesto, cualquier comparación directa, realizada de acuerdo con métodos normalizados, es una forma de obtener trazabilidad. Como ejemplos de métodos normalizados se puede recurrir a las normas publicadas, de las que ISO 5725 es probablemente la más completa, la más seguida y la primera que se citaría en la mayoría de cursos como éste.

A falta de un conjunto de laboratorios dispuestos a compararse con nosotros en el momento que más nos convenga, debemos recurrir a lo que en magnitudes físicas se llamaría la comparación con un patrón. Allí, a esta operación, bien sea directa o bien se haga por intermedio de patrones intermedios, se le llama la calibración de los sistemas de medida, siempre que se haga en condiciones bien establecidas.

Nosotros sabemos perfectamente que para obtener condiciones de medida bien establecidas, hay que normalizar los métodos. Reténganse las dos nociones contenidas en esta expresión: métodos normalizados para comparar resultados con un patrón. En análisis químico la palabra calibración designa una operación diferente, y la obtención de nuestra trazabilidad consiste en validar nuestros métodos de medida normalizados.

No tendría sentido pedir que se modifiquen los nombres, lo que sí importa es señalar la identidad fundamental de unas operaciones y otras que, por otra parte, tienen equivalentes claros en cada uno de los dos campos: análisis químicos y medidas de magnitudes físicas. Lo que tal vez se necesita es abandonar determinados prejuicios. Un patrón no puede definirse ni utilizarse sin la ayuda de métodos de medida normaliza-

dos: deberíamos acostumbrarnos a incorporar ambas cosas indisolublemente y no asustarnos de llamar patrones también a dichos métodos.

Y, por supuesto, llamar también patrones a los materiales de referencia obtenidos y validados mediante tales métodos y al contrario: no aceptar como fuentes válidas de trazabilidad ningún material (muestra de referencia o reactivo de alta pureza), por prestigioso que sea su origen, que no pueda ser validado mediante la comparación de los resultados obtenidos por su mediación.

Por último, y por coherencia, debe haber quedado bien claro que no se puede alegar trazabilidad para nuestros resultados si en la normalización de métodos no se incluye el manejo de la incertidumbre de las correcciones que hay que efectuar a causa de la calibración instrumental (incluyendo desde el simple material volumétrico al instrumento más sofisticado).

En una calibración, en el sentido dado a esa palabra en metrología, se comparan entre sí dos sistemas de generar resultados de medida, uno de los cuales puede considerarse "fiel", porque sus resultados son "verdaderos".

(Un sistema "fiel" es equivalente a un patrón porque sus resultados se distribuyen alrededor del valor del patrón).

En una calibración, en el sentido más peculiar aunque no exclusivo del análisis químico, la magnitud "leída" en un instrumento (con una escala de valores, no una salida única) no es la magnitud de interés y debe ser "traducida" a ésta a través de la aplicación del sistema a valores conocidos del mensurando.

Esto es en realidad más general que la calibración que se ha citado en primer lugar, y que no es más que el caso particular de esta segunda clase de calibraciones, en el que la magnitud leída es también nominalmente la magnitud de interés. La prueba es que la curva de calibración (o la recta, si se elige la regresión lineal) puede representarse de la misma manera en ambos casos.

El caso de la calibración de un mensurando con un único valor de salida se reduce a la trivialidad de ajustar la recta que pasa por el punto problema y por el origen. Para añadir importancia a este caso se debe señalar que es el único en el que sabemos considerar las incertidumbres de ambas coordenadas, aun sin olvidar que también el "origen" es incierto.

Establecidos los conceptos generales dentro de los que se va a desarrollar el proceso de medida químico, puede ser útil estudiar las etapas concretas que conducen a su validación.

## 1. VALIDACIÓN DE MÉTODOS ANALÍTICOS

Validar un método de análisis consiste en verificar y documentar su validez, esto es, su adecuación a unos determinados requisitos establecidos por el usuario.

Estos requisitos reciben el nombre de parámetros de calidad ('figures of merit' o 'performance characteristics' en la literatura en lengua inglesa) y pueden ser de tipo estadístico/quimiométrico o de tipo operativo/económico. Entre los primeros figuran los parámetros fundamentales de exactitud y precisión y los secundarios de selectividad, sensi-

bilidad, límites del método o robustez. Validación, por ejemplo según EURACHEM / WELAC (GÜNZLER, 1994) es la observación de diversas propiedades bajo las cuales los modelos de calibración se han desarrollado: exactitud, precisión, incertidumbre, selectividad, especificidad, linealidad, intervalo de trabajo, sensibilidad, límites de decisión, detección y cuantificación, y robustez. (EURACHEM, 1993) Entre los parámetros de tipo operativo/económicos se encuentran la rapidez del análisis, su costo, la inversión, mantenimiento, etc.

No cabe duda que para verificar la calidad de un método el usuario debe definir previamente los parámetros de calidad en los que está interesado. Dado que es prácticamente imposible lograr valores elevados para todos ellos simultáneamente, debe llegarse a un equilibrio o compromiso entre ellos (MASSART, 1989), (KATEMAN, 1993), (VALCÁRCEL, 1993). Es obvio que tanto la exactitud como la precisión son parámetros generalmente muy apreciados, pero el analista debe estar dispuesto a sacrificar hasta cierto punto su valor numérico en aras de la rapidez, la sencillez o el costo, si así se requiere en circunstancias determinadas.

En el presente capítulo vamos a incidir en los conceptos relacionados con los dos parámetros de calidad fundamentales para validar un método: la verificación de la exactitud y el cálculo del valor numérico de la incertidumbre. Como se verá, ambos conceptos están estrechamente ligados y no se puede verificar el primero sin una estimación del segundo ni tiene sentido calcular el segundo sin haber verificado el primero. Para ello se utilizan diversas herramientas auxiliares: entre las más importantes se encuentran las pruebas de significación o tests de hipótesis, tanto los de tipo paramétrico como aquellos que no se basan en distribuciones conocidas y parámetros estadísticos de los datos a analizar, sin olvidar otros conceptos como el de resultados discrepantes o el número de cifras significativas.

Siendo exactitud y incertidumbre los términos utilizados en el vocabulario del análisis químico, las normas que establecen los criterios para el aseguramiento de la calidad, o garantías de calidad, hablan de trazabilidad e incertidumbre. Todos aquellos que nos dedicamos al análisis químico debemos realizar un pequeño esfuerzo de adaptación y conocer el significado y alcance de estos términos divulgados desde hace ya una década.

## 2. EXACTITUD Y TRAZABILIDAD

El concepto de exactitud, según la ISO 3534, se define como el grado de concordancia entre el resultado de un ensayo y el valor de referencia aceptado. La misma norma especifica que este concepto aplicado a un conjunto de resultados implica una combinación de errores aleatorios y una componente de error sistemático.

Hasta ahora se había considerado la exactitud simplemente como la comprobación de la ausencia de errores sistemáticos, pero para ser consecuentes con la definición del término exactitud que es aplicable no solamente en el campo químico sino en todos los campos de la ciencia, también se tienen que incorporar los errores aleatorios. Es decir, exactitud es suma de dos conceptos: veracidad y precisión. Veracidad, según la norma

ISO 3534, es el grado de concordancia entre el valor medio obtenido a partir de una serie de resultados de ensayo y un valor de referencia aceptado. Diremos que un resultado es veraz si se encuentra libre de errores sistemáticos (o más rigurosamente, si los errores sistemáticos cometidos son aceptables). Precisión, según la misma norma ISO 3534, es el grado de concordancia entre ensayos independientes obtenidos bajo unas condiciones estipuladas. Mediante la precisión se evalúan los errores aleatorios. Por lo tanto diremos que un resultado es exacto si es veraz (se encuentra libre de errores sistemáticos) y preciso (los errores aleatorios son aceptables).

Como se verá posteriormente, la veracidad se evalúa utilizando referencias, y estas nos condicionan el grado de trazabilidad de nuestro resultado.

Para evaluar la veracidad, componente fundamental como ya hemos visto de la exactitud, debemos de conocer el valor considerado como verdadero (el valor de referencia) en la muestra a analizar. Por ello es muy necesario el uso de materiales de referencia, la participación en ejercicios interlaboratorios, la utilización de métodos de referencia o la práctica de otras muchas técnicas de diverso valor dentro de la escala de calidad (TAYLOR, 1987), (DUX, 1990), (GARFIELD, 1991), (VALCÁRCEL, 1992).

El proceso de evaluación de la exactitud se resume en los siguientes pasos:

- 1.- Establecimiento del procedimiento concreto de análisis que se va a utilizar con la muestras de concentración de analito desconocida.
- 2.- Obtención del patrón (en el sentido más amplio del término) de referencia, ya sea un material de referencia, la participación en un ejercicio interlaboratorio, una muestra con cierta concentración de analito añadido ('spiked sample'), etc. Esta etapa conlleva la asignación del valor considerado como verdadero.
- 3.- Análisis repetido del patrón seleccionado en la etapa 2 mediante el procedimiento establecido en la etapa 1. Como resultado de esta etapa se obtiene nuestra propia estimación del valor de referencia (considerado verdadero).
- 4.- Comparación estadística de nuestra estimación y el valor de referencia.
- 5.- Si no existen diferencias estadísticamente significativas entre los dos valores comparados en la etapa 4 y además el laboratorio actúa bajo condiciones de aseguramiento de calidad, se puede afirmar que nuestro procedimiento conduce a resultados veraces. Si además las diferentes estimaciones de precisión nos son satisfactorias, se considerará que nuestro procedimiento conduce a resultados exactos y, por tanto, que se considerarán exactos los resultados de analizar las muestras problema con el mismo procedimiento y condiciones que hemos evaluado.

En la práctica, por tanto, no evaluamos la exactitud de un resultado concreto referido a una muestra problema. En realidad evaluamos el hecho que nuestro procedimiento conduce a resultados exactos para ciertas muestras. Si la comprobación resulta positiva, trasladamos la exactitud del procedimiento a las muestras problema.

El concepto de trazabilidad, de hecho, es muy similar al conocido tradicionalmente como exactitud, sin embargo el término consigue soslayar los problemas de tipo filosófico que plantean los términos 'exacto' y 'verdadero'. Proveniente del campo de la metrología, recordando otra vez su definición, se define formalmente como "la propiedad del resultado de una medición que consiste en que se pueda establecer el resultado previsi-

ble de su comparación directa con los patrones apropiados, por lo general, nacionales o internacionales, por medio de una cadena ininterrumpida de comparaciones reales". En otras palabras, el concepto de trazabilidad se relaciona con la propiedad de establecer los vínculos adecuados entre nuestro resultado y la referencia o patrón adecuados. Un resultado por sí solo no tiene mucho valor si no podemos acreditar su conexión (i.e., establecer la traza o trazar) con una referencia acreditada.

En el campo de las medidas físicas, la trazabilidad se consigue mediante el uso de patrones certificados e instrumentos calibrados que se interconectan entre sí hasta llegar al patrón físico fundamental, por ejemplo, el kilogramo, el metro o el segundo. En el análisis químico no podemos seguir el mismo procedimiento dado que intervienen operaciones como la filtración, la precipitación, la elución cromatográfica, etc., en las que no es posible garantizar la conservación de las masas totales en juego y, como consecuencia, el concepto de trazabilidad parece desdibujarse casi totalmente para la magnitud de interés, que es la concentración de un analito en una muestra determinada.

El concepto de trazabilidad se recupera mediante el uso de los materiales de referencia certificados, la participación en los ejercicios interlaboratorios y el empleo de todas aquellas otras técnicas diseñadas al efecto (TAYLOR, 1987). Por tanto, se hace evidente que determinar la trazabilidad a un valor convencionalmente verdadero determinado es la misma operación que evaluar la veracidad de nuestro procedimiento.

### 3. VERIFICACIÓN DE LA TRAZABILIDAD

Dado que en la práctica existen numerosas formas de conseguir trazabilidad, los modos de evaluarla son también muy diversos. En este apartado y en sus diversos sub-apartados se relacionan los más frecuentes, sin pretender ser exhaustivos.

Antes de iniciar cualquier determinación de un analito (que en terminología establecida por la guía ISO 2 (ISO/IEC Guide 2,1996), recibe el nombre de mensurando) en una muestra, todos los instrumentos, aparatos y material utilizado en el procedimiento de análisis deben haber sido calibrados (cada uno de ellos con una frecuencia que se relaciona el plan de calibración establecido). Durante esta etapa se generan toda una serie de correcciones, las debidas tanto a las imperfecciones del instrumento como a las magnitudes que influyen en su funcionamiento (humedad, temperatura,...), que se integran automáticamente en la medida corregida. Así, por ejemplo, si al calibrar un matraz aforado de 100.00 ml se comprueba que el volumen real es de 100.15 ml (debido, por ejemplo, al propio defecto de fabricación y, por otra parte, a los procesos de dilatación y contracción achacables a los cambios de temperatura que sufra el matraz), se ha de tomar el valor corregido en lugar del valor nominal que podría haber sido correcto en un cierto momento. Se ha introducido, por tanto, una corrección de 0.15 ml en el valor del volumen que tendrá su influencia en los cálculos de la medida final.

### 3.1 COMPARACIÓN DE UNA MEDIA DE UN CONJUNTO DE RESULTADOS CON UN VALOR DE REFERENCIA

Supongamos que disponemos de un material de referencia certificado semejante al tipo de muestra problema que debemos analizar. La institución que certifica el material ha asignado a la concentración de analito de interés un valor  $\bar{x}_r$ , que consideramos como valor verdadero.

Siguiendo las etapas establecidas en el apartado anterior, debemos analizar repetidamente, en condiciones de reproducibilidad (resultados individuales obtenidos con el mismo método pero bajo diferentes condiciones: analistas, instrumentos, laboratorios y tiempo) (MASSART, 1989) (MASSART, 1997), el material de referencia utilizando el mismo procedimiento que usaremos posteriormente para analizar muestras de composición desconocida. Se obtiene de esta forma un conjunto de  $n$  medidas repetitivas. La media aritmética  $\bar{x}_R$  de estos resultados, en ausencia de errores sistemáticos, debería coincidir con el valor verdadero  $\bar{x}_r$ , asignado por la organización que proporciona el material de referencia. En realidad, existen los errores aleatorios y, a pesar que estos se compensan cuando el número de repeticiones es muy elevado, existe una diferencia entre  $\bar{x}_r$  y  $\bar{x}_R$ . El problema se reduce a conocer si esta discrepancia es debida a la presencia de algún error sistemático o sólo a errores aleatorios. Para resolverlo se aplica el test estadístico  $t$  de Student.

Si el número  $n$  de repeticiones llevadas a cabo con el material de referencia es menor que 30 y si los resultados obtenidos siguen una distribución normal se calcula el valor del estadístico  $t$  mediante la expresión:

$$t_{cal} = \frac{|\bar{x}_r - \bar{x}_R|}{s_R / \sqrt{n}} \quad [1]$$

derivada directamente del intervalo de confianza para la media aritmética,  $\bar{x}_R = t_{s_R} / \sqrt{n}$ , y se compara su valor absoluto con el valor de  $t$  tabulado para una probabilidad de cometer un error de 1ª clase,  $\alpha$ , y una probabilidad de error de 2ª clase,  $\beta$ , y  $n-1$  grados de libertad, por ejemplo en (BEYER, 1968) o (CETAMA, 1986). Si  $|t_{cal}| < t_{tab}$  se acepta la hipótesis nula, es decir, se acepta que no existen diferencias estadísticamente significativas entre  $\bar{x}_r$  y  $\bar{x}_R$  y, por tanto, que no se han detectado errores sistemáticos al analizar el material de referencia con la metodología utilizada para el nivel de  $\alpha$  y  $\beta$  elegidos. En otras palabras, el procedimiento aplicado es trazable al material de referencia utilizado.

Cuanto mayor sea la desviación típica,  $s_R$ , de los resultados obtenidos, existen más posibilidades de que  $|t_{cal}|$  sea menor que  $t_{tab}$  y, por tanto, que no se detecten errores sistemáticos. Sin embargo, el analista debe ser consciente que en este caso se sacrificaría la precisión del resultado final en aras de una supuesta trazabilidad. Esta forma de proceder sería totalmente incorrecta, de hecho, antes de evaluar la trazabilidad, debemos asegurarnos que el procedimiento empleado se encuentra bajo control estadístico en nues-

tro laboratorio, es decir, que el valor medido fluctúa en el tiempo, de acuerdo con las leyes de probabilidad, con una varianza determinada alrededor de un valor esperado.

Si el número de repeticiones es mayor que 30 se considera que la distribución será normal y se calcula el estadístico  $z$ :

$$z_{cal} = \frac{|\bar{x}_r - \bar{x}_R|}{\sigma_R / \sqrt{n}} \quad [2]$$

y se compara su valor absoluto con el valor de  $z$  tabulado para los valores de  $\alpha$  y  $\beta$  elegidos (BEYER, 1968), (CETAMA, 1986). Si los resultados no siguen una distribución de densidad de probabilidad Gaussiana deberemos comprobar si existen resultados discrepantes, aplicar alguna técnica de transformación de datos o utilizar tests no paramétricos mencionados en el apéndice del presente capítulo. Un valor discrepante, valor aberrante u *outlier* en inglés (BARTNETT, 1994) (ROUSSEEUW, 1987), es un valor que de alguna manera es inconsistente con el resto de observaciones. Un punto puede ser un valor discrepante a causa de errores aleatorios (en este caso no hace falta tomar medidas encaminadas a revisar el método) o porque el punto no pertenece a la misma población que la muestra.

Si como resultado del procedimiento estadístico se determina que el procedimiento aplicado no es trazable al material de referencia utilizado, se invalida el procedimiento de medida en su globalidad. Sabemos que habremos incorporado algún error sistemático en alguna de las etapas del procedimiento de análisis que se ha transmitido al resultado final de tal forma que queda invalidado. Tendremos que localizarlo y subsanarlo: revisando el procedimiento normalizado de trabajo, recalibrando instrumentos y material de medida, comprobando concentraciones de disoluciones utilizadas, etc. antes de intentar cualquier análisis de muestras problema.

El lector ha de ser consciente del uso de un mismo término, calibración, para dos conceptos distintos: calibración de un instrumento se refiere a la comprobación que su respuesta es la adecuada mientras que también se denomina calibración al proceso de relacionar matemáticamente la respuesta de un instrumento con la concentración de analito responsable de dicha respuesta. Aunque existen similitudes conceptuales entre los dos significados, no ha de existir ninguna duda en lo que respecta a las diferencias entre el proceso de calibración de un instrumento o aparato de medida y la calibración en un método de análisis.

### 3.2 COMPARACIÓN DE LAS MEDIAS DE DOS CONJUNTOS DE RESULTADOS. CONJUNTOS DE RESULTADOS INDEPENDIENTES ENTRE SÍ

Hoy en día es frecuente que los materiales de referencia vayan acompañados de sus certificados correspondientes en los que podemos encontrar una indicación de la incertidumbre asociada al valor medio  $\bar{x}_r$ , considerado por el organismo de certificación como la mejor estimación del valor verdadero correspondiente al analito de interés. En otros

casos puede que no exista material de referencia adecuado y nos veamos obligados a validar nuestro procedimiento de análisis (o procedimiento de ensayo en terminología más general de ISO 2) mediante la utilización de un método de referencia. En ambos casos se obtienen dos conjuntos de datos (los datos del material de referencia o los del método de referencia y el conjunto de medidas repetidas realizadas con nuestro procedimiento) con sus valores medios respectivos  $\bar{x}_r$  y  $\bar{x}_R$  y sus desviaciones típicas  $s_r$  y  $s_R$ . La verificación de la trazabilidad de nuestro procedimiento de análisis se lleva a cabo mediante la comparación de las medias aritméticas  $\bar{x}_r$  y  $\bar{x}_R$  de estos dos conjuntos de datos.

Si el número de resultados  $n_r$  o  $n_R$  de los dos conjuntos de datos considerados independientes que se comparan es menor que 30, si cada uno de ellos sigue una función de densidad de probabilidad Gaussiana y si las varianzas  $s_r^2$  y  $s_R^2$  de cada conjunto no difieren de forma estadísticamente significativa, se calcula el valor  $t$  de Student según la ecuación:

$$t_{cal} = \frac{\bar{x}_r - \bar{x}_R}{\sqrt{s^2 (1/n_r + 1/n_R)}} \quad [3]$$

donde  $s^2$  es la varianza conjunta ('pooled variance') de los dos grupos de resultados calculada mediante la expresión:

$$s^2 = \frac{(n_r - 1)s_r^2 + (n_R - 1)s_R^2}{(n_r + n_R - 2)} \quad [4]$$

y se compara el valor de  $|t_{cal}|$  con el valor teórico de  $t_{tab}$  que se encuentra en las tablas para el valor de  $\alpha$  y  $\beta$  elegidos y  $(n_r + n_R - 2)$  grados de libertad, por ejemplo en (BEYER, 1968), (CETAMA, 1986). Si  $|t_{cal}| < t_{tab}$  acepta la hipótesis nula concluyéndose que no existen diferencias estadísticamente significativas entre las medias aritméticas obtenidas con los dos métodos para los valores de probabilidad de falso positivo y falso negativo considerados.

Si el número de resultados  $n_r$  y  $n_R$  son iguales o mayores que 30 se calcula el valor  $z$ :

$$z_{cal} = \frac{|\bar{x}_r - \bar{x}_R|}{\sqrt{\sigma^2 (1/n_r + 1/n_R)}} \quad [5]$$

cuya función de densidad de probabilidad es Gaussiana con valor medio igual a cero y desviación típica la unidad. El valor de  $|z_{cal}|$  se compara con el valor teórico  $z_{tab}$  que se encuentra en tablas estadísticas para  $\alpha$  y  $\beta$  elegidos.

La comparación de las varianzas  $s_r^2$  y  $s_R^2$  de dos conjuntos de datos se lleva a cabo mediante el test  $F$  de Fisher. Se calcula la relación  $F_{cal}$ :

$$F_{cal} = \frac{s_1^2}{s_2^2} \quad [6]$$

de tal forma que aquella de las dos varianzas  $s_r^2$  ó  $s_R^2$  que sea mayor se sitúa en el numerador. Esta relación calculada se compara con el valor tabulado para el nivel de significancia elegido, el riesgo de cometer un error de segundo tipo,  $\beta$ , y los grados de libertad  $(n_r - 1)$  y  $(n_R - 1)$  correspondientes al numerador y al denominador respectivamente. Si  $F_{cal}$  es mayor que  $F_{tab}$  se concluye que existe una diferencia estadísticamente significativa entre las varianzas de los dos conjuntos de datos.

En este último caso no puede calcularse la varianza conjunta  $s^2$  y el test  $t$  para la comparación entre las dos medias de los dos conjuntos de resultados, llamado test de Cochran, se lleva a cabo calculando el estadístico:

$$t_{cal} = \frac{|\bar{x}_r - \bar{x}_R|}{\sqrt{(s_r^2/n_r) + (s_R^2/n_R)}} \quad [7]$$

y comparando  $|t_{cal}|$  con  $t_{tab}$  para los valores de probabilidad de falso positivo y falso negativo considerados y los grados de libertad resultantes de redondear al número entero más próximo el valor:

$$\text{grados de libertad} = \frac{(s_r^2/n_r + s_R^2/n_R)^2}{\frac{(s_r^2/n_r)^2}{1} + \frac{(s_R^2/n_R)^2}{1}} \quad [8]$$

### 3.3 COMPARACIÓN DE DOS CONJUNTOS DE RESULTADOS. CONJUNTOS DE RESULTADOS RELACIONADOS O PAREADOS

La comparación de las medias de dos conjuntos de datos mediante el test  $t$  presupone que se verifica la ausencia/presencia de errores sistemáticos para una única concentración de analito. No se comprueba si los posibles errores presentes dependen de la concentración de analito medida. Para obviar este hecho, normalmente se lleva a cabo la experimentación preparando un conjunto de muestras en las que la concentración de analito varía en el intervalo de valores más frecuentes que van a encontrarse en la práctica, y se analizan con los dos métodos que se pretenden comparar.

En el caso que interese comparar, por ejemplo, dos laboratorios o dos métodos diferentes, incorporando la influencia que la concentración pueda tener en la detección de errores sistemáticos, se analizan por cada método el mismo conjunto de muestras, que contienen distintas concentraciones de analito. Los conjuntos de datos reciben el nom-

bre de relacionados o pareados dado que existe un valor hallado con cada método para cada concentración de analito. En estos casos, se halla la diferencia  $d_i$  para cada par de valores correspondientes, y se calcula la media aritmética  $\bar{x}_d$  y la desviación típica  $s_d$  del conjunto de diferencias. Si el número de diferencias  $d_i$  es igual o mayor que 30, se calcula el valor  $z$

$$z_{cal} = \frac{\bar{x}_d}{s_d / \sqrt{n}} \quad [9]$$

y se compara con los valores tabulados  $z_{tab}$  que siguen la distribución Gaussiana para un nivel de significancia y una probabilidad de falso negativo dados.

Si el número de diferencias  $d_i$  es menor que 30, la expresión:

$$t_{cal} = \frac{\bar{x}_d}{\sigma_d / \sqrt{n}} \quad [10]$$

sigue una distribución  $t$  de Student con lo que se compara  $|t_{cal}|$  con  $t_{tab}$  con  $(n_i - 1)$  grados de libertad. Extrayendo las mismas conclusiones que en casos anteriores.

### 3.4 COMPARACIÓN DE DOS CONJUNTOS DE RESULTADOS. APLICACIÓN DE LA REGRESIÓN LINEAL

Una alternativa a la técnica de validación descrita en el apartado anterior la constituye la aplicación de la regresión lineal por mínimos cuadrados. Este es el caso común en que los resultados obtenidos mediante un nuevo método se comparan con aquellos obtenidos mediante un método bien establecido.

La ausencia de todo error, ya sea aleatorio o sistemático, en los datos se manifestaría mediante la obtención de una línea recta de pendiente unidad y ordenada en el origen cero, tal como muestra la línea (a) de la Figura 1. La presencia de sólo un error sistemático proporcional llevaría a la obtención de una recta, representada en la Figura 1 por la línea (b), con pendiente distinta a la unidad, mientras que la presencia de un error sistemático constante conduciría a la obtención de una recta con una ordenada en el origen distinta a cero (Figura 1(c)). Los errores aleatorios, que acompañan siempre a todo tipo de resultados, darían lugar a una dispersión de los puntos experimentales alrededor de la línea de regresión. De este modo, la presencia de los tres tipos de errores mencionados, aleatorios, sistemáticos constantes y sistemáticos proporcionales darían lugar a la recta (d) de la Figura 1.

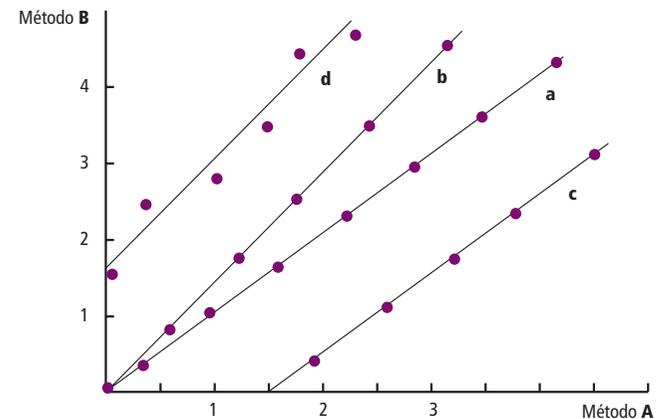


Figura 1. Comparación mediante regresión lineal de resultados obtenidos mediante dos métodos analíticos distintos

El análisis de regresión parece el más indicado para detectar y cuantificar este tipo de errores (MILLER, 1991),(DRAPER, 1981). La desviación estándar de los residuos de la regresión u otros parámetros relacionados (VANKEERBERGHEN, 1992) proporcionan una medida para cuantificar los errores aleatorios. Los tests independientes de la pendiente y de la ordenada en el origen permiten comprobar la ausencia de errores sistemáticos debidos al blanco o a la presencia de errores proporcionales (BOQUÉ, 1994). Sin embargo hay que señalar que, desde el punto de vista estadístico, existe una diferencia notable cuando la técnica de regresión lineal se aplica a la calibración o cuando se utiliza, como en esta ocasión, para comparar dos conjuntos de resultados obtenidos experimentalmente. Este hecho es debido a que en este último caso ninguno de los dos conjuntos de datos está libre de error, por lo que se tendrían que aplicar técnicas de regresión que contemplen la presencia de errores en ambos ejes de coordenadas  $x$  e  $y$  (RIU, 1995), a diferencia de la calibración en donde se considera que sólo el eje  $x$  está libre de error. En todo caso, si se desea utilizar la técnica usual de regresión por mínimos cuadrados, debido principalmente a su simplicidad, en el eje de las  $x$  se colocará el método analítico que produzca un menor error.

Para llevar a cabo la comparación de dos métodos analíticos a múltiples niveles de concentración se puede utilizar el test conjunto de la ordenada en el origen y la pendiente. Este test, (MANDEL, 1957) aplicado desde 1957 a problemas químicos, busca la recta de regresión entre los resultados obtenidos al analizar diversas muestras a varios niveles de concentración mediante dos métodos, método de referencia, normalmente en el eje de las  $x$ , y método a comprobar, normalmente en el eje de las  $y$ . Si los dos métodos en comparación producen resultados que no difieren estadísticamente entre ellos a un

nivel de significancia  $\alpha$ , la ordenada de la recta de regresión no ha de ser estadísticamente diferente de 0 y simultáneamente la pendiente de la recta de regresión no ha de ser estadísticamente diferente de 1. El principal inconveniente de este test es que para buscar los coeficientes de la recta de regresión utiliza el método de mínimos cuadrados, el cual considera el método situado en el eje de las  $x$  como libre de error. Recientemente se ha desarrollado el test conjunto para la ordenada en el origen y la pendiente considerando errores en los dos métodos analíticos (RIU 1996) el cual permite llevar a cabo la comparación de dos métodos analíticos teniendo en cuenta los errores reales asociados a cada uno de los métodos en comparación.

### 3.5 COMPARACIÓN DE LAS MEDIAS DE DIVERSOS CONJUNTOS DE RESULTADOS. ANÁLISIS DE LA VARIANZA (ANOVA)

La comparación de medias correspondientes a diversos conjuntos de resultados puede ser habitual en los laboratorios analíticos. Este es el caso que nos encontramos en los ejercicios interlaboratorios, si bien en esta ocasión no somos nosotros sino el equipo promotor del ejercicio el que aplica el ANOVA para verificar si alguno de los laboratorios participantes proporciona resultados que se diferencian de forma estadísticamente significativa del resto. Otro ejemplo de aplicación en nuestro laboratorio podría ser cuando se quieren comparar entre sí diversos métodos con distintas características, o con un método de referencia o cuando la comparación quiere establecerse entre varios analistas que analizan una muestra con el mismo método. En definitiva, ANOVA se utiliza para *"analizar medidas que dependen de varios tipos de efectos que operan simultáneamente con el doble fin de decidir cuales de ellos son importantes y de poder estimarlos"* (SCHEFFÉ, 1953).

El nombre "análisis de la varianza" puede conducir a error sobre la aplicación de este test. Claramente, ANOVA se utiliza para comparar medias de diversos conjuntos de resultados, no sus varianzas. El nombre de esta técnica proviene del hecho que utiliza la comparación de parámetros estadísticos, en forma de varianzas, para llegar a la conclusión sobre las medias en estudio.

En el presente apartado vamos a referirnos a la comparación de diversos conjuntos de resultados que difieren en sólo un factor o causa de variación. Si, por ejemplo, se desean comparar  $k$  laboratorios que determinan  $n_j$  veces la concentración de un analito en una misma muestra con un mismo método, el objetivo consiste en detectar si alguno de los laboratorios proporciona resultados que difieren de forma estadísticamente significativa de los demás.

En el ANOVA se han de cumplir tres tipos de hipótesis, aunque se aceptan ligeras desviaciones respecto a las condiciones ideales:

- 1.- Cada conjunto de datos es independiente de los demás.
- 2.- Los resultados obtenidos para cada conjunto siguen una distribución normal.
- 3.- Las varianzas de cada conjunto de datos no difieren de forma significativa.

Siguiendo el ejemplo propuesto, si se obtienen  $n_j$  resultados para cada laboratorio, se dispone de un conjunto de datos que se puede ordenar de la forma dispuesta en la Tabla 1.

Laboratorio	Resultados		Media aritmética
	1	2 3 ... i ... j ... n	
1	$X_{11}$	$X_{12} X_{13} \dots X_{1i} \dots X_{1j} \dots X_{1n}$	$\bar{X}_1$
2	$X_{21}$	$X_{22} X_{23} \dots X_{2i} \dots X_{2j} \dots X_{2n}$	$\bar{X}_2$
3	$X_{31}$	$X_{32} X_{33} \dots X_{3i} \dots X_{3j} \dots X_{3n}$	$\bar{X}_3$
...	...	...	...
$i$	$X_{i1}$	$X_{i2} X_{i3} \dots X_{ii} \dots X_{ij} \dots X_{in}$	$\bar{X}_i$
...	...	...	...
$k$	$X_{k1}$	$X_{k2} X_{k3} \dots X_{ki} \dots X_{kj} \dots X_{kn}$	$\bar{X}_k$

Tabla 1. Disposición de resultados en el Análisis de la Varianza

En dicha Tabla,  $x_{ij}$  sería el resultado  $j$  del laboratorio  $i$ . El objetivo del análisis de la varianza consiste en la comparación de las distintas medias aritméticas  $\bar{x}_i$  ( $i = 1, 2, \dots, k$ ) con el fin de determinar si alguna de ellas difiere significativamente de las demás. Para ello utiliza una estrategia muy lógica: si los resultados proporcionados por los distintos laboratorios no contienen errores sistemáticos, las medias aritméticas respectivas no diferirán mucho unas de otras y su dispersión, debida a los errores aleatorios siempre presentes, será comparable a la dispersión "tipo" cometida individualmente por cada laboratorio.

Se hallan, por tanto, dos varianzas como medidas de las dispersiones comentadas y se comparan mediante el conocido test- $F$ . La utilización de estas varianzas son las que dan nombre a la técnica, ANOVA.

El resultado  $x_{ij}$  va acompañado de un error  $\varepsilon_{ij}$  que refleja la diferencia entre este valor y el valor considerado 'verdadero'  $\mu$ :

$$\varepsilon_{ij} = x_{ij} - \mu \tag{11}$$

dicho error puede descomponerse del modo siguiente:

$$\varepsilon_{ij} = (x_{ij} - \bar{x}_i) + (\bar{x}_i - \bar{x}) + (\bar{x} - \mu) \tag{12}$$

donde

$$\bar{x}_j = \frac{\left( \sum_{j=1}^{n_j} x_{ij} \right)}{n_j} \quad [13]$$

$$\bar{x} = \frac{\left( \sum_{i=1}^k \sum_{j=1}^{n_j} x_{ij} \right)}{N} \quad [14]$$

$$N = \sum_{i=1}^k n_j \quad [15]$$

siendo  $n_j$  el número de resultados repetidos que proporciona el laboratorio  $i$  ( $x_{ij} - \bar{x}_j$ ) describe la desviación dentro del laboratorio  $i$  ( $x_i - \bar{x}$ ) describe la desviación del laboratorio  $i$  respecto a la media aritmética de todos los laboratorios y, finalmente,  $(\bar{x} - \mu)$  describe una desviación sistemática general que no puede calcularse si no se conoce el valor considerado como verdadero  $\mu$ . Normalmente se considera  $x = \mu$  y, por tanto,  $(\bar{x} - \mu) = 0$ . En estas condiciones se puede reescribir:

$$\varepsilon_{ij} = (x_{ij} - \bar{x}_i) + (\bar{x}_i - \bar{x}) \quad [16]$$

$$\varepsilon_{ij} = e_{ij} + \delta_i \quad [17]$$

donde  $e_{ij}$  = error residual o residual y  $\varepsilon_i$  = error entre laboratorios o laboratory bias. Dado que:

$$\varepsilon_{ij} = x_{ij} - \mu \quad \text{o} \quad \varepsilon_{ij} = x_{ij} - \bar{x} \quad [18]$$

se deduce que:

$$x_{ij} = \bar{x} + e_{ij} + \delta_i \quad [19]$$

ecuación lineal que expresa cada medida  $x_{ij}$  como suma de una constante  $\bar{x}$  y de dos variables aleatorias,  $e_{ij}$  que varía dentro del laboratorio y  $\varepsilon_i$  que varía entre laboratorios.

El análisis de la varianza consiste en calcular una serie de sumas de cuadrados que conducirán a las varianzas finales:

$$\varepsilon_{ij} = (x_{ij} - \bar{x}) = (x_{ij} - \bar{x}_i) + (\bar{x}_i - \bar{x}) \quad [20]$$

$$(x_{ij} - \bar{x})^2 = (x_{ij} - \bar{x}_i)^2 + (\bar{x}_i - \bar{x})^2 + 2(x_{ij} - \bar{x}_i)(\bar{x}_i - \bar{x}) \quad [21]$$

y sumando para todas las  $i$  y todas las  $j$ :

$$\begin{aligned} \sum_{i=1}^k \sum_{j=1}^{n_j} (x_{ij} - \bar{x})^2 &= \sum_i \sum_j (x_{ij} - \bar{x}_i)^2 + \sum_i n_j (\bar{x}_i - \bar{x})^2 + \\ &+ 2 \sum_j (x_{ij} - \bar{x}_i) \sum_i (\bar{x}_i - \bar{x}) \end{aligned} \quad [22]$$

siendo el último término de la ecuación igual a cero, dado que la suma de desviaciones con respecto a las medias aritméticas es siempre cero. La suma de cuadrados total  $SS_T$  puede escribirse como la suma de dos sumas de cuadrados:

$$\sum_i \sum_j (x_{ij} - \bar{x})^2 = \sum_i \sum_j (x_{ij} - \bar{x}_i)^2 + \sum_i n_i (\bar{x}_i - \bar{x})^2 \quad [23]$$

$SS_R$  mide las desviaciones entre observaciones,  $x_{ij}$ , de cada laboratorio y la media del grupo, por tanto es una medida de la dispersión dentro de los laboratorios. Al dividir  $SS_R$  por los correspondientes grados de libertad,  $N-k$  se obtiene la varianza "dentro de los laboratorios".  $SS_{lab}$  mide las desviaciones entre laboratorios y, dividida por sus grados de libertad,  $k-1$ , constituye la varianza "entre laboratorios".

Si no existe diferencia estadísticamente significativa entre estas dos últimas, la presencia de errores aleatorios será la causa predominante de la discrepancia de valores medios. Si, por el contrario, existe algún error sistemático,  $SS_{lab}/(k-1)$  será mucho mayor que  $SS_R/(N-k)$  con lo que el valor calculado  $F$

$$F_{cal} = \frac{SS_{lab}/(k-1)}{SS_R/(N-k)} \quad [24]$$

será mayor que el valor tabulado  $F_{tab}$  para la probabilidad  $\alpha$ , la probabilidad  $\beta$  elegidas, y los grados de libertad mencionados. En este último caso se podría concluir que, con el grado de probabilidad escogido, al menos uno de los laboratorios produce resultados cuyo valor medio difiere de forma estadísticamente significativa de los demás.

El análisis de la varianza no indica cuantos laboratorios difieren ni cuales son. Una inspección visual de los resultados sin duda puede proporcionar pistas, sin embargo, con el fin de tener criterios más sólidos, existen distintos tests para conocer aquellos laboratorios cuyos resultados discrepan significativamente del resto (WINER, 1970).

De la misma forma puede plantearse la detección de errores sistemáticos en los resultados proporcionados por distintos analistas utilizando cada uno, por ejemplo, dos instrumentos distintos comunes a todos ellos. Si se llevan a cabo repeticiones de los análisis efectuados con cada instrumento, ANOVA proporciona información sobre la existencia de discrepancias entre analistas y entre instrumentos. Este sería un ejemplo de un ANOVA de dos factores (MASSART, 1989), (MASSART, 1997), (MEIER, 1993).

## 4 PRECISIÓN E INCERTIDUMBRE

Todo resultado analítico se expresa de modo general como:

$$\text{Valor 1} \pm \text{Valor 2} \quad [25]$$

donde Valor 1 representa la mejor estimación que podemos proporcionar de la concentración de analito presente en la muestra analizada. Valor 1 está relacionado con la trazabilidad o la exactitud tal como hemos desarrollado en los apartados precedentes. El Valor 2 se relaciona con la incertidumbre del resultado final, es decir, con la indicación del intervalo en el que podremos encontrar, con un margen de seguridad determinado, el Valor 1. La incertidumbre está relacionada, por tanto, con la dispersión de resultados que obtendríamos al repetir el análisis completo de la muestra problema un número determinado de veces, aparte de englobar otros aspectos.

Incertidumbre y precisión de un resultado analítico son términos muy relacionados. Quizás aquellos que nos dedicamos al análisis químico estamos acostumbrados a asociar el término precisión a un determinado múltiplo de la desviación típica o a un intervalo de confianza resultante de repetir el análisis de la muestra problema. El término incertidumbre quiere ser más globalizador, en el sentido de considerar todas las fuentes posibles de error que intervienen el resultado final. Pero la diferencia más importante se encuentra en el hecho que el concepto de incertidumbre está íntimamente ligado con el concepto de trazabilidad, no así la precisión. Así la ISO (ISO, 1993) define la incertidumbre como "una estimación unida al resultado de un ensayo que caracteriza el rango de valores dentro de los cuales se afirma que está el valor verdadero". Por lo tanto, la ISO nos está diciendo que ya que la estimación de la incertidumbre debe incluir el valor verdadero, no tiene ningún sentido calcular la incertidumbre sin antes haber verificado la trazabilidad de nuestro procedimiento analítico, ya que solamente en estas circunstancias podremos asegurar que el valor verdadero cae dentro de nuestra incertidumbre. (ANALY-

TICAL METHODS COMMITTEE, 1995) Precisión, en cambio, indica una mera dispersión de los datos.

Sabemos que el procedimiento de medida químico se compone de diversas etapas, normalmente independientes entre sí. Esta es la característica fundamental que diferencia las medidas de tipo químico de aquellas de tipo físico. Al repetir diversas veces todo el conjunto de etapas, el teorema del límite central proporciona el soporte teórico que permite considerar que los resultados obtenidos se distribuyen generalmente de acuerdo a curvas Gaussianas. Este hecho es muy importante para que podamos realizar todo el conjunto de inferencias basadas en la estadística paramétrica.

Es evidente que cada una de las etapas del procedimiento analítico contribuye en cierta medida a la incertidumbre final. Este hecho adquiere mayor relevancia cuando se pretende optimizar la precisión del método, debiendo incidir en aquellas etapas que más incidencia poseen en la incertidumbre del resultado final.

Por otro lado, existen diversos tipos de componentes que inciden en la incertidumbre global. Nosotros estamos acostumbrados a trabajar en aquellas que se estiman mediante procedimientos estadísticos (denominadas de Tipo A). Existen, tal como desarrollaremos más adelante, otras componentes que se aprecian por otros métodos y reciben el nombre de componentes Tipo B.

Debemos observar que la verificación de la trazabilidad también puede considerarse una etapa en el procedimiento global de obtención del resultado final. En esta etapa también existen errores y, por tanto, de la misma manera tendremos que evaluar su contribución en la incertidumbre final.

Dada la gran variedad de metodologías analíticas disponibles, existen diversas formas de calcular la incertidumbre del resultado final. (ELLISON, 1997) En el presente capítulo vamos a suponer, en primer lugar, que estamos interesados en calcular la incertidumbre de resultados en un laboratorio de control de calidad.

En este tipo de laboratorios no es fácil que se posean las condiciones para repetir un cierto número de veces el análisis de cada una de las muestras problema. Rigurosamente hablando, si no se repiten análisis no es posible el cálculo de cualquier medida de la dispersión. Por otra parte, si se obliga a repetir análisis de cada muestra problema, es probable que muchos laboratorios no puedan cumplir con el requisito de proporcionar la incertidumbre. Una solución de compromiso para este problema viene dada al seguir una estrategia parecida a la utilizada para evaluar la trazabilidad. Podemos intentar calcular una *incertidumbre asociada al procedimiento* al analizar repetidamente una muestra tipo parecida a las muestras problema. Estando en condiciones de aseguramiento de calidad (i.e. cumpliendo los requisitos de las normas de la serie EN 45.000) podremos *trasladar esta incertidumbre del procedimiento a la incertidumbre del resultado final*.

Si nuestro laboratorio está en condiciones de analizar repetidamente cada una de las muestras problema, caso tratado en segundo lugar, es obvio que la estimación de la precisión en estas condiciones es mucho mejor dado que calculamos *la incertidumbre asociada a la muestra que formará parte de la incertidumbre del resultado final*. Volvemos, sin embargo, al balance expuesto al inicio del capítulo, rapidez, tiempo, costo o productividad frente a una buena estimación, en este caso, de la incertidumbre. Cada usuario debe valorar convenientemente su situación.

A continuación vamos a exponer dos métodos para calcular la incertidumbre asociada al proceso de medida: el método auspiciado por la ISO, comúnmente llamado *bottom-up*, y un método basado en el trabajo realizado en el proceso de verificación de la trazabilidad del procedimiento analítico.

## 5 CÁLCULO DE LA INCERTIDUMBRE

### 5.1 CÁLCULO DE LA INCERTIDUMBRE MEDIANTE EL MÉTODO DE LA ISO

El método para el cálculo de incertidumbres promovido por la ISO, y llamado comúnmente *bottom-up*, (ISO,1999 y 2005) fue primeramente aplicado a resultados físicos y posteriormente adaptado al campo químico por el EURACHEM. (CITAC/EURACHEM 2000)

Este método recibe el nombre de *bottom-up* debido a que divide el proceso de medida químico en sus partes fundamentales, calcula la incertidumbre de cada una de las partes y las combina para obtener la incertidumbre global del proceso de medida químico. Esta aproximación está basada en estimar la incertidumbre total mediante la identificación, estimación y combinación de todas las fuentes de incertidumbre asociadas con el proceso de medida: se divide el proceso de medida en bloques o pasos, se busca la incertidumbre de cada bloque o paso, y finalmente se combinan para encontrar la incertidumbre total. El procedimiento para cuantificar la incertidumbre total se muestra en el esquema siguiente:



#### 5.1.1 ESPECIFICACIÓN

El resultado de la medida depende de diversos parámetros como son otros mensurandos que puedan haberse estimado, constantes y otras cantidades. La relación entre la cantidad medida ( $y$ ) y estos parámetros ( $x_1, x_2, x_3, \dots$ ) debe ser escrita, siempre que sea posible, como una expresión algebraica del tipo  $y = f(x_1, x_2, x_3, \dots)$ .

Supongamos que queremos determinar la concentración de ácido clorhídrico a partir de una valoración ácido-base. La concentración estimada depende, en su forma más simple, de los siguientes parámetros: la concentración de hidróxido sódico y los volúmenes de ácido y de base. La relación entre ellos es la ecuación

$$N_{HCl} = \frac{N_{NaOH} \cdot V_{NaOH}}{V_{HCl}}$$

Donde  $N$  representa la normalidad y  $V$  el volumen en litros de ácido y de base. Si la concentración de NaOH está estandarizada con ftalato ácido de potasio,  $N_{NaOH}$  depende a su vez del peso de reactivo, de su pureza, de su peso equivalente y del volumen equivalente de la solución de NaOH.

Esta aproximación puede ser usada con métodos clásicos de análisis o bien con otros casos muy simples. Ya se ha comentado que esta aproximación tuvo sus orígenes en las medidas de tipo físico. Para otras metodologías, las dificultades para establecer una relación matemática entre los resultados analíticos y los parámetros que intervienen el procedimiento analítico son evidentes. Por tanto, normalmente será mucho mejor dividir el proceso de medida en un número de bloques como el muestreo, los pretratamientos de la muestra, la separación del analito, la medida instrumental, los cálculos ..., en lugar de intentar buscar una expresión algebraica del tipo  $y = f(x_1, x_2, x_3, \dots)$ . Una vez hecho esto, se llevará a cabo la identificación y la cuantificación de las fuentes de incertidumbre de cada bloque. Posteriormente, la incertidumbre total será calculada combinando los valores de las incertidumbres individuales calculadas.

Siguiendo con el ejemplo de la valoración, algunos de los bloques en que podría dividirse la determinación de ácido clorhídrico son la preparación de la disolución de NaOH, la medida de una alícuota de NaOH, la valoración de NaOH con HCl y el cálculo de la concentración de la disolución de HCl.

#### 5.1.2 IDENTIFICACIÓN

Para cada uno de los parámetros o bloques de los que depende el mensurando, las fuentes de incertidumbre asociadas con éstos deben ser identificadas. El objetivo de este apartado es identificar todas las fuentes de incertidumbre existentes independientemente de la importancia que tengan cada una de ellas.

Sin pretender ser exhaustivos, las fuentes posibles de incertidumbre en un procedimiento analítico son el muestreo, los pretratamientos de la muestra, el analista, la

corrección de posibles errores sistemáticos del instrumento, la calibración, la pureza de los reactivos, las condiciones de medida, la contaminación y los errores computacionales.

En el ejemplo de la valoración del HCl, todas las fuentes de incertidumbre de cada bloque deben ser identificadas. Por ejemplo, en la medida de la alícuota de NaOH alguna de las fuentes de incertidumbre identificadas son la variación al llenar la pipeta hasta la marca (errores aleatorios), el volumen nominal de la pipeta (debido a su calibración) y la temperatura (ya que la pipeta podría ser usada a una temperatura diferente de la de calibración). En el bloque correspondiente a la valoración, la presencia de interferencias (otra sustancia con propiedades ácidas), la inestabilidad del NaOH (ya que se carbonata con CO<sub>2</sub>) y la detección del punto final son algunas de las fuentes de incertidumbre.

### 5.1.3 CUANTIFICACIÓN

Las fuentes de incertidumbre identificadas en el apartado anterior deben ser cuantificadas. Como no todos los componentes identificados tendrán una contribución significativa en la incertidumbre total, debe hacerse una selección preliminar de los componentes que contribuirán de forma importante basada en sus incertidumbres individuales.

La cuantificación de los componentes individuales puede hacerse de tres formas diferentes:

- " a través del trabajo experimental, frecuentemente mediante el análisis de réplicas.
- " usando datos y resultados obtenidos en trabajos experimentales previos.
- " usando cualquier tipo de información disponible sobre la variabilidad del parámetro evaluado y frecuentemente asumiendo un valor basado en la experiencia del analista.

La cuantificación basada en un análisis estadístico de los resultados experimentales dan lugar a una evaluación de la incertidumbre de tipo A. La repetibilidad, reproducibilidad u otras estimaciones de precisión intermedias se usan frecuentemente para evaluar la importancia de los efectos aleatorios.

La evaluación experimental es la mejor forma de estimar los componentes de incertidumbre. Sin embargo, hay veces que los experimentos no pueden realizarse debido a falta de tiempo, a la no disponibilidad de los recursos necesarios, o a la dificultad intrínseca en la estimación de la incertidumbre de un determinado bloque o parámetro. En este caso, la incertidumbre debe evaluarse a partir de trabajos anteriores o bien a partir del criterio del analista. Estas estimaciones están definidas como una evaluación de la incertidumbre de tipo B.

Las estimaciones basadas en la utilización de datos anteriores se obtienen a partir de la información dada por los proveedores: tolerancia del material volumétrico ... Estudios anteriores pueden proporcionar información de la incertidumbre asignada a los valores de materiales de referencia o bien de los errores sistemáticos de los métodos detectados a través de ejercicios de intercomparación.

Sin embargo, en otros casos la única forma de estimar la incertidumbre está basada en la opinión del analista. Por ejemplo, la evaluación de las diferencias entre el nivel de concentración del analito en el material de referencia y en la muestra, la cuantificación de la dependencia de la concentración del analito en estudios de recuperación o la evaluación de las diferencias de la composición de la matriz entre el material de referencia y la muestra analizada se estiman a partir de los conocimientos que tiene el analista sobre el mensurando y el procedimiento de medida. Estas estimaciones no son completamente objetivas. Para limitar esta desventaja, deben estar basadas en un conocimiento detallado de la naturaleza del mensurando y del método y procedimiento de medida usado.

*Cálculo de las incertidumbres estándar de cada bloque o parámetro.* Para obtener la incertidumbre total combinada del proceso de medida siguiendo el principio de propagación de los errores (MELOUN, 1994), la distribución de valores originados a partir de efectos aleatorios y sistemáticos tienen que estar expresados en las mismas unidades: las incertidumbres estándar. Por tanto, puede ser necesario la conversión de otras medidas de dispersión a incertidumbres estándar.

Cuando las incertidumbres individuales se evalúan experimentalmente (estimaciones de tipo A), la desviación estándar calculada a partir de los datos es igual al término de incertidumbre estándar. Por tanto, no es necesaria la conversión de otras medidas de dispersión.

Supongamos que se quiere calcular la incertidumbre estándar debida a una dilución realizada con un matraz aforado. La información del matraz dada por el proveedor es  $250 \pm 0.15$  ml. Si tenemos motivos para asumir una distribución Gaussiana con un nivel de confianza del 95%, entonces la incertidumbre estándar se calcula dividiendo 0.15 ml por 1.96 (el valor de  $z$  tabulado para ese nivel de confianza).

Si dispusiéramos de más información, por ejemplo, si la distribución del valor fuera una distribución  $t$ , 0.15 debería dividirse por el valor de  $t$  tabulado para el número de grados de libertad considerados.

Sin embargo, muchas veces no hay información sobre la función de distribución ni sobre el número de grados de libertad. En estos casos, únicamente puede asumirse que es igual de probable que el valor esté en cualquier lugar dentro de los límites del intervalo. Por tanto, debe asumirse una distribución rectangular de los valores. En este caso la incertidumbre estándar se obtiene dividiendo 0.15 por  $\sqrt{3}$ .

Cuando las estimaciones de tipo B están basadas en el juicio del analista, se asume que la probabilidad asignada está basada en el sentido común y en el conocimiento y experiencia del analista. En este caso, es posible estimar directamente el componente de incertidumbre como desviación estándar. Si esto no es posible, debe hacerse una estimación acerca de la función de distribución. Una distribución rectangular debe asumirse cuando no hay motivos para creer que los errores pequeños son más probables que los grandes dentro de los límites del intervalo. Si los errores pequeños son más probables, la desviación estándar puede calcularse dividiendo el valor correspondiente a la mitad del intervalo por 3 o por  $\sqrt{6}$  (dependiendo de que el analista asuma una distribución normal o bien una distribución triangular). (ISO, 1993)

### 5.1.4 COMBINACIÓN

Los componentes individuales de incertidumbre estimados como incertidumbres estándar, deben combinarse siguiendo el principio de propagación de los errores para dar lugar a la incertidumbre combinada. Si una cantidad medida,  $y$ , es función de muchos parámetros,  $y = f(x_1, x_2, x_3, \dots)$ , la incertidumbre combinada estándar viene dada por la expresión general:

$$u(y(x_{1,j}, \dots)) = \sqrt{\sum_{i=1,n} \left( \frac{\partial y}{\partial x_i} \cdot u(x_i) \right)^2 + \sum_{\substack{i,k=1,n \\ i \neq k}} \left( \frac{\partial y}{\partial x_i} \cdot \frac{\partial y}{\partial x_j} \cdot \text{COV}(x_{ik}) \right)} \quad [26]$$

donde:

$x_i$  ( $i = 1, \dots, n$ ) son los parámetros de los que depende el resultado del mensurando  $u(x_i)$  incertidumbre del parámetro  $i$  expresada como incertidumbre estándar  $\text{COV}(x_{ik})$  es la covarianza entre  $x_i$  y  $x_k$

Si todos los parámetros que contribuyen al resultado del mensurando son independientes, el término correspondiente a la covarianza es cero. Por tanto, la expresión general usada para calcular la incertidumbre combinada estándar se simplifica:

$$u(y(x_{1,j}, \dots)) = \sqrt{\sum_{i=1,n} \left( \frac{\partial y}{\partial x_i} \cdot u(x_i) \right)^2} \quad [27]$$

Si la relación que hay entre el mensurando y los parámetros se expresa únicamente a través de sumas y diferencias, por ejemplo  $y = x_1 + x_2 - x_3 + \dots$ , la incertidumbre estándar combinada viene dada por:

$$u(y(x_{1,j}, \dots)) = \sqrt{\sum_{i=1,n} (u(x_i))^2} = \sqrt{u(x_1)^2 + u(x_2)^2 + u(x_3)^2 + \dots} \quad [28]$$

Esta expresión indica que la incertidumbre combinada estándar es siempre más grande que el mayor componente individual de incertidumbre estándar. Al mismo tiempo, puede verse que la contribución de los términos pequeños puede ser despreciada la mayoría de las veces.

**Cálculo de la incertidumbre combinada expandida.** La incertidumbre expandida debe calcularse para proporcionar un intervalo que incluya una elevada proporción de todos los valores que pueden atribuirse al mensurando.

La incertidumbre expandida proporciona un intervalo correspondiente a un nivel de confianza especificado y se obtiene multiplicando la incertidumbre combinada estándar por el factor de incertidumbre,  $k$ . Generalmente,  $k=2$  o  $k=3$  pueden utilizarse para obtener aproximadamente un nivel de confianza del 95% o del 99% respectivamente. En la práctica,  $k=2$  es el valor recomendado en la mayoría de las ocasiones.

### 5.2 CÁLCULO DE LA INCERTIDUMBRE BASADO EN UTILIZAR LA INFORMACIÓN OBTENIDA EN EL PROCESO DE VERIFICACIÓN DE LA TRAZABILIDAD

Esta aproximación (MAROTO, 1999a), (MAROTO, 1999b) está basada en calcular la incertidumbre aprovechando la información obtenida en la verificación de la trazabilidad de un método analítico. Como la trazabilidad debe ser siempre verificada, esta aproximación no implica un gran trabajo adicional para calcular la incertidumbre. Siguiendo la ley de propagación de errores, el componente de incertidumbre obtenido a partir de la verificación de la trazabilidad se combina con otros componentes no considerados previamente (muestreo, submuestreo, etapas previas no tenidas en cuenta en la verificación de la trazabilidad...) para obtener la incertidumbre total asociada al resultado analítico.

En la verificación de la trazabilidad se genera información sobre la varianza asociada al método analítico en condiciones intermedias de tiempo,  $s_R^2$ . Recordemos que  $s_R^2$  viene dada por la expresión:

$$s_R^2 = \frac{\sum_i (x_i - \bar{x})^2}{n - 1} \quad [29]$$

donde cada  $x_i$  representa el resultado de un análisis de la muestra de referencia en *condiciones intermedias de tiempo* dentro del laboratorio (o más estrictamente, condiciones intermedias de precisión del laboratorio) y  $n$  es el número de análisis realizados en dichas condiciones. Las condiciones intermedias de tiempo, es decir, cambio de analista, instrumento, patrones, ubicación, condiciones de uso y tiempo, garantizan la obtención de una estimación de la incertidumbre asociada a los resultados de las muestras problema que incorpora la variación de estos factores. Nos interesa tener esta estimación porque los resultados de una misma muestra problema obtenidos en diferentes condiciones (por ejemplo en días diferentes y/o por diferentes analistas) deben ser comparables. Es decir, el resultado obtenido unos días después debe estar contenido dentro del intervalo de valores proporcionado por el resultado del primer análisis realizado y la incertidumbre asociada al método analítico.

Los componentes de incertidumbre asociados a un resultado analítico, que deben ser estimados y posteriormente combinados, se agrupan en cuatro términos. Esto queda ilustrado en la Figura 2:

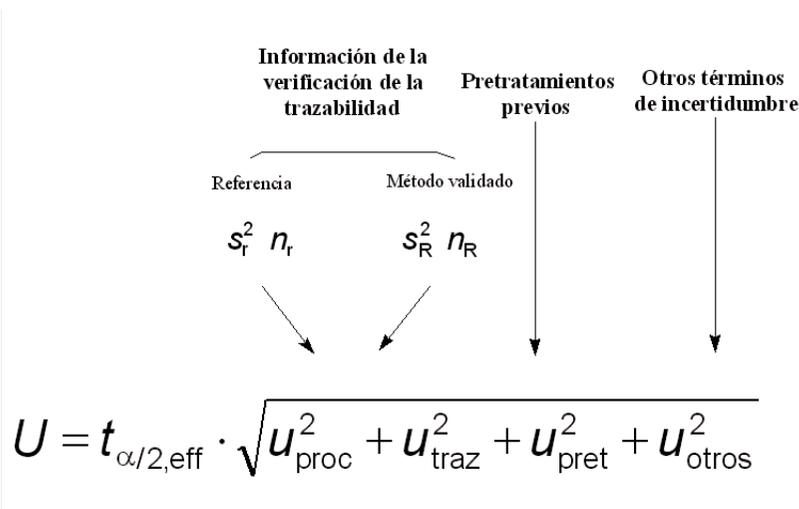


Figura 2. Cálculo de la incertidumbre total utilizando la información obtenida en el proceso de verificación de la trazabilidad.

### 5.2.1. INCERTIDUMBRE ASOCIADA AL PROCEDIMIENTO ANALÍTICO

Esta incertidumbre es debida a la variación experimental obtenida cuando se aplica el procedimiento analítico a muestras futuras. Este término está relacionado con la variabilidad de los resultados analíticos debida a los errores aleatorios que se producen en las diferentes etapas del procedimiento analítico. Si cada una de las muestras futuras se analizase un número elevado de veces,  $n_s$ , en condiciones intermedias de tiempo, se podría evaluar este término de incertidumbre como:

$$u_{\text{proc}} = \frac{s_R}{\sqrt{n_s}} \quad [30]$$

Sin embargo, la mayoría de veces el resultado de una muestra futura se obtiene a partir de un único análisis ( $n_s=1$ ) o como mucho a partir de tres análisis ( $n_s=3$ ). En este caso, la varianza en condiciones intermedias del método no puede obtenerse a partir de los resultados de los análisis de la muestra futura ya que el número de análisis realizados es insuficiente. Para solucionar este problema, lo que se hace es utilizar la varianza obtenida en condiciones intermedias de tiempo estimada en el proceso de la verificación de la trazabilidad del método. Esto puede hacerse siempre que la muestra futura analizada sea similar a la utilizada en el proceso de verificación de la trazabilidad y siempre que las condiciones de aseguramiento de la calidad se lleven a cabo en el laboratorio de forma

efectiva. Es importante puntualizar que mediante el uso de diseños experimentales (por ejemplo se puede consultar la guía ISO 5725-3) como los diseños anidados (*nested designs*), se puede descomponer  $s_R$  en sus componentes de repetibilidad ( $s_r$ ) i entre series ( $s_{\text{series}}$ ), con lo que si una futura muestra de rutina es analizada por ejemplo dos veces en condiciones de repetibilidad, la información obtenida al efectuar este duplicado puede ser utilizada para disminuir el término correspondiente a la repetibilidad. En cambio, si solo se dispone de  $s_{\text{p}}$  no se puede aprovechar el trabajo efectuado en el duplicado ya que  $s_R$  se refiere a condiciones intermedias de tiempo.

### 5.2.2. INCERTIDUMBRE DE LA VERIFICACIÓN DE LA TRAZABILIDAD

En la verificación de la trazabilidad del método analítico se comprueba si el error sistemático (sesgo, o bias en inglés) del laboratorio y/o del método es estadísticamente significativo. El tipo de sesgo depende de la referencia usada al comprobar la trazabilidad del método (si se utiliza un método de referencia dentro del mismo laboratorio se comprueba únicamente si el sesgo del método es significativo. Con un material de referencia se verificaría conjuntamente si el sesgo del método y del laboratorio son o no significativos).

Si el sesgo no es significativo, el método es trazable a la referencia utilizada pero, sin embargo, hay un componente de incertidumbre debida al propio proceso de verificación de la trazabilidad. El sesgo,  $\delta$ , se estima a partir de los resultados generados en la verificación de la trazabilidad como:

$$\delta = \bar{x}_R - \bar{x}_r \quad [31]$$

donde  $\bar{x}_R$  es la media de los resultados analíticos de la muestra de referencia obtenidos en condiciones intermedias de tiempo y  $\bar{x}_r$  es el valor asignado a la muestra de referencia. La incertidumbre asociada al propio proceso de verificación de trazabilidad viene dada por la incertidumbre del sesgo estimado. Esta incertidumbre viene dada por una desviación estándar que llamaremos  $s_d$ :

$$u_{\text{traz}} = s_d \quad [32]$$

Esta desviación estándar se estima a partir de la varianza del método trazado y de la varianza asociada al valor asignado a la muestra de referencia. Si ambas varianzas no son significativamente diferentes,  $s_d$  se obtiene a partir de la varianza conjunta o "pooled variance":

$$s_d = \sqrt{\left( \frac{(n_R - 1) \cdot s_R^2 + (n_r - 1) \cdot s_r^2}{n_R + n_r - 2} \right) \left( \frac{1}{n_R} + \frac{1}{n_r} \right)} \quad [33]$$

donde  $s_r$  es la desviación estándar asociada al valor de referencia y  $n_r$  es el número de resultados utilizados para obtener el valor de referencia. Los grados de libertad asociados a  $s_r$  son  $(n_R + n_r - 2)$ .

Si las varianzas son significativamente diferentes,  $s_d$  se calcula como:

$$s_d = \sqrt{\frac{s_R^2}{n_R} + \frac{s_r^2}{n_r}} \quad [34]$$

Los grados de libertad se calculan con la ecuación 8 (aproximación de Satterthwaite (SATTERTHWAITE, 1941)).

### 5.2.3. INCERTIDUMBRE ASOCIADA A ETAPAS PREVIAS

Este componente es necesario para incluir las fuentes de incertidumbre asociadas al submuestreo o bien a etapas previas (filtrado, pesado, secado, etc.) no consideradas en el proceso de verificación de la trazabilidad.

Esta incertidumbre puede ser estimada a partir de una muestra que tenga las mismas características que las muestras futuras que se analizarán posteriormente. El submuestreo y/o etapas previas tienen que hacerse variando representativamente todos los factores de variabilidad que los pueden afectar (operador, material, tiempo, etc.). Los resultados analíticos de cada muestra submuestreada y/o pretratada se obtienen bajo condiciones de repetibilidad (mismo analista, día, calibrado, etc.). Esto se hace para no sobreestimar la incertidumbre ya que, si las muestras se analizaran bajo condiciones de intermedias de tiempo, la variabilidad de los resultados no correspondería únicamente a la variabilidad experimental de las etapas previas y/o del submuestreo sino que también correspondería a la variabilidad experimental del método. Por tanto, estaríamos incluyendo un componente de incertidumbre que ya ha sido considerado previamente (apartado 5.2.1).

La incertidumbre de este término corresponde a la desviación estándar de los resultados analíticos de las muestras submuestreadas y/o pretratadas,  $s_{pret}$

$$u_{pret} = s_{pret} \quad [35]$$

El número de grados de libertad asociados a  $u_{pret}$  es  $n_{pret}-1$  (siendo  $n_{pret}$  el número de muestras submuestreadas y/o pretratadas). En este término también queda incluida la repetibilidad del método analítico. Como la repetibilidad del método suele ser despreciable frente a las condiciones intermedias de tiempo debida a etapas previas y/o submuestreo, la incertidumbre total apenas se sobreestima. En todo caso, si se dispone de la desviación estándar en condiciones de repetibilidad, ésta se puede sustraer de  $s_{pret}$ .

### 5.2.4. INCERTIDUMBRE DEBIDA A OTRAS FUENTES DE VARIACIÓN

Este último término contiene todos los otros componentes de incertidumbre que no han sido considerados previamente. Muchos de estos componentes están considerados como incertidumbres de "Tipo B". Algunos ejemplos de fuentes de incertidumbre incluidos en este término son debidas a diferencias entre la muestra de referencia utilizada en la verificación de trazabilidad y la muestra futura (o muestra de trabajo) o bien asociada al sesgo de la referencia utilizada en la verificación de la trazabilidad (se estima trazando esta referencia a otra con una mayor nivel de trazabilidad). La estimación de estos términos no es fácil. Suele evaluarse a partir de la experiencia y conocimiento del analista y también a partir de otro tipo de información (certificados de calibración, resultados de estudios de intercomparación...).

Todas los términos de incertidumbre debidos a fuentes de variación no considerados previamente, se combinan siguiendo la ley de propagación de errores:

$$u_{otros} = \sqrt{\sum_i s_i^2} \quad [36]$$

donde  $s_i$  es la incertidumbre estándar asociada a cada una de las fuentes de variación no consideradas previamente. Como es una varianza combinada, el número de grados de libertad se obtiene a partir de la aproximación de Satterthwaite:

$$v_{otros} = \frac{(u_{otros}^2)^2}{\sum_i \frac{(s_i^2)^2}{v_i}} \quad [37]$$

donde  $v_{otros}$  son los grados de libertad asociados a  $u_{otros}$  y  $v_i$  los grados de libertad asociados a  $s_i$ . Sin embargo, muchas veces no se dispone de información sobre los grados de libertad. En este caso, se utiliza el factor de incertidumbre,  $k$  (ya explicado en la aproximación propuesta por la ISO).

### 5.2.5. ESTIMACIÓN DE LA INCERTIDUMBRE TOTAL

Una vez han sido calculadas las varianzas de cada término y sus grados de libertad asociados, la incertidumbre total estándar se calcula combinando estas varianzas.

$$u = \sqrt{u_{proc}^2 + u_{traz}^2 + u_{pret}^2 + u_{otros}^2} \quad [38]$$

La incertidumbre total expandida se obtiene multiplicando la incertidumbre total estándar por el valor de  $t$  efectivo tabulado de dos colas a una probabilidad  $\alpha$  y con el número de grados de libertad efectivos,  $v_{\text{eff}}$  obtenidos a partir de la aproximación de Satterthwaite [37].

$$U = t_{\alpha/2, \text{eff}} \cdot \sqrt{u_{\text{proc}}^2 + u_{\text{traz}}^2 + u_{\text{pret}}^2 + u_{\text{otros}}^2} \quad [39]$$

La incertidumbre total expandida proporciona un intervalo de confianza en el cual el valor de referencia está incluido con una probabilidad  $1-\alpha$ . En vez de utilizar el valor de  $t$  efectivo, puede usarse el factor de cobertura,  $k$  para obtener la incertidumbre total expandida. No es una mala aproximación siempre y cuando el número de grados de libertad de las varianzas combinadas sea elevado (especialmente el de la varianza que tenga una contribución mayor). De todas formas, es más correcto utilizar el valor de  $t$  efectivo.

Para obtener una buena estimación de la incertidumbre total es necesario que las muestras futuras sean similares a la utilizada en la verificación de la trazabilidad y a la utilizada para estimar la incertidumbre de los pretratamientos y/o el submuestreo (en nivel de concentración y matriz) y que además el aseguramiento de la calidad esté implantado de forma efectiva en el laboratorio.

## 6 APÉNDICE

### 6.1 PRUEBAS DE SIGNIFICACIÓN

Las pruebas de significación o tests de hipótesis son las herramientas que poseemos para responder objetivamente a las cuestiones estadísticas que se plantean a lo largo del proceso analítico.

El analista, en el proceso de validación está interesado en el problema fundamental de la trazabilidad del método empleado. La cuestión: ¿es este método trazable? se reformula en términos de pruebas de significación, por ejemplo, en: ¿existe una diferencia estadísticamente significativa entre nuestro valor medio y el valor considerado como verdadero? Las pruebas de significación también se utilizan para comprobar distintas hipótesis que se consideran tales como la uniformidad de varianzas, la presencia o ausencia de valores discrepantes o la adecuación del conjunto de datos a distribuciones determinadas.

Una hipótesis consiste en la suposición que se realiza sobre un determinado suceso; por ejemplo, que la mínima concentración medible de analito con un determinado procedimiento difiera de forma estadísticamente significativa del blanco analítico, mientras que un test estadístico es la regla que permite dilucidar la validez de dicha hipótesis.

Planteadas una hipótesis,  $H$ , se pueden adoptar dos decisiones: aceptarla o rechazarla. Así mismo puede ser que dicha hipótesis sea cierta o falsa. Los resultados se esquematizan en el siguiente cuadro:

#### Decisión adoptada mediante el test

		aceptación	rechazo
Situación	$H_0$ cierta	correcto	error 1ª especie
	$H_0$ falsa	error 2ª especie	correcto

en donde puede observarse que la decisión resulta correcta al aceptar una hipótesis que es cierta o rechazar una hipótesis falsa. Si se rechaza una hipótesis cierta se comete un error de tipo  $\alpha$  o de primera especie mientras que si se acepta una hipótesis falsa se comete un error tipo  $\beta$  o de segunda especie.

Las consecuencias de ambos errores suelen ser muy distintas. En nuestro interés por la trazabilidad, el error de 1ª especie o falso positivo implica que existe una diferencia estadísticamente significativa entre las medias cuando en realidad no la hay, en otras

palabras, que nuestro método no es trazable cuando en realidad lo es. Por el contrario, la comisión del error de 2ª especie o falso negativo, implica que consideramos al método trazable cuando en realidad no lo es; es decir, cometeremos errores sistemáticos en el análisis de las muestras problema. En este caso, como ocurre generalmente, el rechazar una hipótesis que sea cierta suele tener mayor trascendencia que aceptar una hipótesis que sea falsa. Los analistas, hasta el momento, no hemos concedido la suficiente importancia a los errores tipo  $\beta$  cuando en realidad vemos que pueden ser más importantes que los errores tipo  $\alpha$ . A la cantidad  $(1 - \beta)$  se le denomina potencia del test. En realidad sabemos que es la probabilidad de rechazar la hipótesis nula cuando en realidad es falsa. La potencia del test depende de  $\alpha$ , del número de observaciones realizadas y de la diferencia real entre los valores comparados.

Para llevar a cabo un test estadístico, en primer lugar es necesario elegir la hipótesis. En estadística existen de dos tipos:

1. la hipótesis nula,  $H_0$ , que establece que no hay diferencias entre valores, es decir, que las diferencias observadas se deben a los errores aleatorios y
2. la hipótesis alternativa,  $H_1$ , que establece la presencia de diferencias entre valores de los parámetros medidos.

En segundo término, se elige el test estadístico y se calcula el parámetro estadístico de interés (mediante las denominadas funciones discriminantes del test o modelo estadístico), nos referimos a  $t_{cal}$ ,  $F_{cal}$  o el adecuado para el test paramétrico o no paramétrico que realizamos.

En tercer lugar se ha de elegir el nivel de significancia,  $\alpha$ , que queremos asignar al test estadístico, teniendo en cuenta también el error tipo  $\beta$ .

Definida la región crítica como el conjunto de valores, normalmente extremos, tales que cuando  $H_0$  es cierto, la probabilidad de que observemos alguno de ellos es realmente muy baja (la probabilidad es  $\alpha$ ).

Finalmente, se observa si el valor calculado se encuentra dentro de la región crítica, en este caso se rechaza la hipótesis nula. Si el valor calculado no se encuentra dentro de la región crítica la decisión es que  $H_0$  no puede rechazarse (que sutilmente no es lo mismo que  $H_0$  se cumple forzosamente). Normalmente este proceso se lleva a cabo comparando el valor calculado con el mismo parámetro tabulado en las tablas estadísticas y se acepta o rechaza la hipótesis elegida. Alternativamente, se puede calcular el nivel de significancia,  $\alpha$ , y la probabilidad  $\beta$  con la que se igualan los parámetros estadísticos calculado y tabulado.

Supongamos que analizando un material de referencia, cuya concentración de analito de interés tiene asignada un valor medio de:  $\bar{x}_r = 5.00$  ppm, se llevan a cabo 12 determinaciones repetidas en condiciones de intermedias de tiempo mediante el procedimiento utilizado en nuestro laboratorio, obteniéndose un valor medio de  $\bar{x}_R = 4.95$  ppm y una desviación estándar,  $s_R = 0.07$  ppm. Bajo la suposición que infinitas repeticiones como las realizadas darían lugar a unas medidas distribuidas según una función de densidad de probabilidad normal, (las 12 medidas siguen una distribución  $t$  con 11 grados de libertad).

La hipótesis nula establece que no existe diferencia estadísticamente significativa entre  $\bar{x}_r = 5.00$  ppm y  $\bar{x}_R = 4.95$  ppm. Para comprobarla se calcula  $t_{cal}$  en nuestro caso,  $t_{cal} = 2.37$ . Dado que  $t_{tab} = 2.201$ , para un nivel de significancia,  $\alpha = 0.05$  y 11 grados de libertad, es mayor que  $t_{cal}$  se rechaza la hipótesis nula y se concluye que nuestro método no es trazable.

Para alcanzar esta conclusión se ha tenido en cuenta el error de 1ª especie con un valor de  $\alpha = 0.05$ . Este nivel de significancia, se define como la probabilidad de rechazar la hipótesis nula cuando en realidad es cierta. Supongamos que en lugar de  $\alpha = 0.05$  se considera el valor  $\alpha = 0.01$ . Ello significa que reducimos la probabilidad de cometer un falso positivo, en otras palabras, reducimos la probabilidad de concluir que las dos medias son distintas cuando en realidad proceden de la misma población estadística. En términos más analíticos, reducimos la probabilidad de concluir que no somos trazables cuando en realidad lo somos. Para  $\alpha = 0.01$ ,  $t_{cal} = 3.106$  con lo que concluiríamos ahora que, con este nivel de significancia, somos trazables (ya que ahora  $t_{cal} < t_{tab}$ ).

Vemos pues que cuanto menor es  $\alpha$ , más probabilidades tenemos de aceptar la hipótesis nula. Sin embargo, este aparente beneficio conlleva la contrapartida del error de 2ª clase. En otras palabras, debemos considerar cual es la probabilidad que tenemos de aceptar que somos trazables cuando en realidad no lo somos.

Si consultamos ahora a la tabla de distribución  $t$  en la que se consideran  $\alpha$  y  $\beta$  (BEYER, 1968), (CETAMA, 1986) observamos que con 11 grados de libertad y  $\alpha = 0.05$ , la probabilidad de cometer un error de 2ª especie es aproximadamente de  $\beta = 0.35$ . Si  $\alpha$  disminuye a 0.01,  $\beta$  aumenta hasta aproximadamente 0.6, es decir, un 60% de probabilidades de cometer un, importante por sus consecuencias, falso negativo.

Para un conjunto de resultados determinado, al aumentar  $\alpha$  disminuye  $\beta$  y viceversa. La única forma de disminuir tanto  $\alpha$  como  $\beta$  es aumentando el número de análisis realizados. La relación entre  $\alpha$ ,  $\beta$  y el número de observaciones para cada test estadístico también se representa en ábacos denominados a menudo curvas características (BEYER, 1968), (CETAMA, 1986).

En las recopilaciones de tablas estadísticas y en los apéndices de libros de estadística se encuentran dos tipos de tablas distintas para cada uno de los valores de  $\alpha$  y  $\beta$ . Se denominan tablas de una cola o de dos colas. La elección de la tabla correspondiente del test que se efectúa depende de la formulación inicial de la hipótesis alternativa,  $H_1$ , que se plantea. Así, por ejemplo, se puede plantear que: "Existe diferencia entre un valor considerado como verdadero,  $\bar{x}_r$ , y la media aritmética,  $\bar{x}_R$ , de un conjunto de resultados", o, de forma alternativa que: "la media de un conjunto de resultados  $\bar{x}_R$  es mayor que un cierto valor dado  $\bar{x}_r$ ". En el primer caso, el valor de la media aritmética  $\bar{x}_R$  puede ser diferente, en términos estadísticamente significativos, tanto por ser mayor como por ser menor que el valor considerado como verdadero  $\bar{x}_r$  con el que se compara; por tanto, los valores  $\alpha$  y  $\beta$  están repartidos entre los dos extremos de la distribución Gaussiana, y el test sería de dos colas.

Sin embargo, cuando se plantea la hipótesis de que la media aritmética  $\bar{x}_R$  es o no mayor que el valor de referencia  $\bar{x}_r$ , la información requerida es de calidad inferior a la del caso anterior. En esta situación, la probabilidad de rechazar la hipótesis nula en caso de ser cierta,  $\alpha$ , o la de aceptar esta hipótesis en caso de ser falsa,  $\beta$ , están concentradas en el extremo superior de la distribución Gaussiana y el test estadístico recibe el nombre

de test de una cola. Debe observarse que tanto si la región crítica está situada en ambos extremos de la distribución o en uno solo, la probabilidad  $\alpha$  es la misma en ambos casos aunque, naturalmente, los valores tabulados del parámetro correspondiente al test estadístico no lo sean.

Si se rechaza  $H_0$  sabemos que  $\bar{x}_R > \bar{x}_r$  de forma estadísticamente significativa. Sin embargo, si se acepta la hipótesis nula se concluye que  $\bar{x}_R$  no es mayor que  $\bar{x}_r$ , y no sabemos, en este último caso, si  $\bar{x}_R$  es igual a  $\bar{x}_r$  o si  $\bar{x}_R < \bar{x}_r$ , en términos estadísticos.

## 7 BIBLIOGRAFÍA

- **Analytical Methods Committee, Analyst, 120, 1995, 2303-2308**
- **AOAC International, AOAC Peer-Verified Methods, Policies and Procedures, Arlington, 1993**
- **AOAC International - ISO 17025 and the Laboratory - An Introduction to Laboratory. Accreditation: 2000**
- **AOAC International - Quality Assurance Principles for Analytical Laboratories - 3rd Edition 2000 - F M Garfield, E Klesten, J Husch ISBN-0-935584-70-6**
- **V. Barnett y T. Lewis, Outliers in Statistical Data, 3rd ed., John Wiley & Sons, Chichester, 1994**
- **W.H. Beyer, Handbook of Tables for Probability and Statistics, 2ª Ed., CRC Press, Boca Raton, 1968.**
- **R. Boqué y F.X. Rius, Profundizando en la Calibración Lineal Univariante en R. Cela (Eds) Avances en Quimiometría Práctica, Servicio de Publicaciones de la Universidad de Santiago de Compostela, Santiago de Compostela, 1994.**
- **E.P. Box, W.G. Hunter y J.S. Hunter, Statistics for Experimenters, an Introduction to Design, Data Analysis and Model Building, Wiley, Nueva York, 1978.**
- **CETAMA, Statistique appliquée a l'exploitation des mesures, 2ª Ed., Masson, París, 1986.**
- **CITAC/EURACHEM, Quality Assurance for Research and Development and Non-Routine Analysis: 1998**
- **CITAC/EURACHEM, Quantifying Uncertainty in Analytical Measurement: 2000**
- **CITAC/EURACHEM Guide on Traceability in Chemical Measurements: 2002**
- **N. Draper, H. Smith, Applied Regression Analysis, 2ª Ed., Wiley, Nueva York, 1981.**
- **J.P. Dux, Handbook of Quality Assurance for the Analytical Chemistry Laboratory, 2ª Ed., Van Nostrand Reinhold, Nueva York, 1990.**
- **S. Ellison, W. Wegscheider y A. Williams, Anal. Chem. News & Features, 1995, 607A-613A**
- **EURACHEM/WELAC Guide 1, Accreditation of Chemical Laboratories, Laboratory of the Government Chemist, Londres 1993**
- **EURACHEM, The Fitness for Purpose of Analytical Methods: A Laboratory Guide to Method Validation and Related Topics: 1998**
- **F.M. Garfield, Quality Assurance principles for Analytical laboratories, 2ª Ed., AOAC, Arlington, 1991.**
- **H. Günzler (Ed.), Accreditation and quality assurance in analytical chemistry, Springer-Verlag, Berlin, 1994**
- **G.J. Hahn y W.Q. Meeker, Statistical Intervals. A Guide for Practitioners, John Wiley & Sons, Nueva York, 1991**
- **Gy, Pierre, Sampling for analytical purposes, Chichester: Wiley, 1998**

- **ISO, International Organization for Standardization, ISO Guide 33:** Uses of Certified Reference Materials for the Assessment of a Measurement Procedure, **Ginebra, 1984.**
- **ISO, International Organization for Standardization, ISO 3534-1:** Statistics Vocabulary and Symbols, **Ginebra, 1993**
- **ISO/BIPM/IEC/IFCC/IUPAC/IUPAP/OIML,** International vocabulary of basic and general terms in metrology (VIM) - **2nd edition 1993**
- **ISO, International Organization for Standardization, ISO 5725-3:** Accuracy (trueness and precision) of measurement methods and results, **Ginebra, 1994**
- **ISO,** Guide to the Expression of Uncertainty in Measurement (GUM), **Geneva Switzerland, 1995.**
- **ISO/IEC Guide 2: 1996** Standardization and related activities - General vocabulary (**currently under revision as ISO 17000**)
- **ISO/IEC 17025:1999 y 2005** General requirements for the competence of testing and calibration Laboratories
- **IUPAC, International Union of Pure and Applied Chemistry (2002)** Recommendations for the use of the term "recovery" in analytical procedures, <<http://www.iupac.org/>>
- **G. Kateman y F.W. Pijpers,** Quality Control in Analytical Chemistry, **Wiley, Nueva York, 1981.**
- **G. Kateman y L. Buydens,** Quality Control in Analytical Chemistry, **2ª Ed., Wiley, Nueva York, 1993.**
- **C. Liteanu y I. Rica,** Statistical Theory and Methodology of Trace Analysis, **Ellis Horwood, Nueva York, 1980.**
- **J. Mandel y F.J. Linnig,** Anal. Chem., **1957, 29, 743-749**
- **A. Maroto, J. Riu, R. Boqué y F.X. Rius,** Estimating uncertainties of analytical results using informations from the validation process, Anal. Chim. Acta, **1999, 391, 173-185**
- **A. Maroto, R. Boqué, J. Riu y F.X. Rius,** Evaluating uncertainty in routine analysis, Trends Anal. Chem., **1999, 18, 577-584**
- **D.L. Massart, B.G.M. Vandeginste, S.N. Deming, Y. Michotte y L. Kaufman,** Chemometrics: a Textbook, **Elsevier, Amsterdam, 1989.**
- **D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S. de Jong, P.J. Lewi y J. Smeyers-Verbeke,** Handbook of Chemometrics and Qualimetrics: Part A, **Elsevier, Amsterdam, 1997.**
- **P.C. Meier y R.E. Zünd,** Statistical Methods in Analytical Chemistry, **Wiley, Nueva York, 1993.**
- **M. Meloun, J. Militký y M. Forina,** Chemometrics for Analytical Chemistry, Vol. I, **Ellis Horwood, Londres, 1994**
- **J.C. Miller y J.N. Miller,** Statistics for Analytical Chemistry, **4ª Ed., Ellis Horwood, Chichester, 1998**
- **J.N. Miller,** Basic Statistical methods for Analytical Chemistry. Part 2. Calibration and Regression Methods. A Review, Analyst, **1991, 116, 3-14.**

- **J.C. Olucha y F.X. Rius,** Software to test the hypothesis of normal distribution of data according to Kolmogorov's test, Trends Anal. Chem., **1990, 9, 77-79.**
- **G. Ramis y M.C. García Alvarez-Coque en Quimiometría, M. Blanco y V. Cerdà, Eds., Publicacions de la Universitat Autònoma de Barcelona, Bellaterra, 1988.**
- **J. Riu, F.X. Rius,** Univariate Regression Models with Errors in Both Axes, **J. Chemometrics, 1995, 9, 343-362.**
- **J. Riu, F.X. Rius,** Assessing the Accuracy of Analytical Methods Using Linear Regression with Errors in Both Axes, Anal. Chem., **1996, 68, 1851-1857**
- **P.J. Rousseeuw y A.M. Leroy,** Robust Regression and Outlier Detection, **Wiley, Nueva York, 1987.**
- **F.E. Satterthwaite,** Psychometrika, **1941, 6, 309-316.**
- **H. Scheffé,** Analysis of Variance, **Wiley, Nueva York, 1953.**
- **W.A. Shewhart,** Economic Control of the Quality of Manufactured Products Van Nostrand, **Nueva York, 1931.**
- **S. Siegel,** Nonparametric Statistics for the Behavioral Sciences, **McGraw-Hill, Nueva York, 1956.**
- **J.K. Taylor,** Quality Assurance of Chemical Measurements, **Lewis Publ., Michigan, 1987.**
- **P. Vankeerberghen y J. Smeyers-Verbeke,** The quality coefficient as a tool in decisions about the quality of calibration in graphite furnace atomic absorption spectrometry. Chemomet. Intel. Lab. Systems, **1992, 15, 195-202.**
- **M. Varcárcel y A. Ríos (Eds),** La Calidad en los Laboratorios Analíticos, **Reverté, Barcelona, 1992.**
- **M. Varcárcel y A. Ríos,** The hierarchy and relationships of analytical properties, Anal. Chem., **1993, 65, 781A-787A.**
- **G.T. Wernimont, W. Spendley,** Use of statistics to develop and evaluate analytical methods, **AOAC, Washington, 1985**
- **B.J. Winer,** Statistical Principles in Experimental Design, **McGraw-Hill, Londres, 1970.**

## BIBLIOGRAFÍA SEGÚN CONCEPTOS Y PARÁMETROS ANALÍTICOS DE CALIDAD

### ACCURACY

- **A framework for in-house accuracy validation of analytical procedures.** A.M. García-Campana et al. Biomedical Chromatography **2000, 14, 27-29**

### COLLABORATIVE STUDIES

- **AOAC International. Guidelines for Collaborative Study Procedures to Validate Characteristics of a Method of Analysis.** J. AOAC Int., **1995, 78 (5), 143A-168A**
- **W. Horwitz. Protocol for the Design, Conduct and Interpretation of Method Performance Studies.** Pure Appl. Chem., **1995, 67, 331-343**

#### INTERNAL QUALITY CONTROL

- M. Thompson, R. Wood. Harmonized Guidelines for Internal Quality Control in Analytical Chemistry Laboratories. Pure Appl. Chem., 1995, 67, 49-56
- European Commission, Quality Control procedures for Pesticides Residues Analysis - Guidelines for Residues Monitoring in the European Union, 2nd Edition, Document n°. SANCO/3103/2000, Brussels, 2000

#### METROLOGY IN CHEMISTRY

- Metrology in Chemistry, why?, what? How? Proceedings of the Maribor Conference. Maribor, 20-21 September 2001. pp 1-8 (comparability, concordance, equivalence, target value for uncertainty,...)

#### PERFORMANCE CHARACTERISTICS

- "AOAC Peer-Verified methods program, Manual on Policies and procedures" AOAC International, Gaithersburg, MD, 1993

#### PROFICIENCY AND METHOD PERFORMANCE STUDIES

- Thompson, M. Wood, R., (1993) The International Harmonised Protocol for the Proficiency Testing of (Chemical) Analytical Laboratories. Pure Appl. Chem., 65, 2123-2144. Also published in J. AOAC Intl. 76, 926-940.
- International Measurement Evaluation Programme (IMEP). Proceedings of the Maribor Conference. Maribor, 20-21 September 2001. Pp 35-40
- Horwitz, W., (1995) Protocol for the Design, Conduct and Interpretation of Method- Performance Studies. Pure Appl. Chem., 67, 331-343
- ICH (1994) Harmonised Tripartite Guideline prepared within the Third International Conference on Harmonisation of Technical Requirements for the Registration of Pharmaceuticals for Human Use (ICH), Text on Validation of Analytical Procedures, 1994, <<http://www.ifpma.org/ich1.html>>

#### RECOVERY

- M. Thompson, S.L.R. Ellison, A. Fajgelj, P. Willetts, R. Wood. Harmonised guidelines for the use of recovery information in analytical measurements, Pure & Appl. Chem., 1999, 71, 337-348
- D.T. Burns, K. Danzer, A. Townshend, Recommendations for the Use of the Term "Recovery" in Analytical Procedures. IUPAC Recommendations 2001 <<http://www.iupac.....>>
- Quality control procedures for pesticide residue analysis, Guidelines for residues monitoring in the European Union, Document 7826/VI/97, European Commission, Brussels, 1997

- IUPAC, International Union of Pure and Applied Chemistry (1999b) Harmonised guidelines for the use of recovery information in analytical measurement, Pure and Applied Chemistry 71, 337-348

#### ROBUSTNES

- Youden procedure. Statistical manual of the AOAC
- Vander-Heyden, Y., Nijhuis A., Smeyers-Verbeke, J., Vandeginste, B.G.M., Massart, D.L., (2001) Journal of Pharmaceutical and Biomedical Analysis 24, 723-753.

#### SELECTIVITY

- J. Vessman, R.I. Stefan, J.F. van Staden, K. Danzer, W. Lindner, D. T. Burns, A. Fejgelj, H. Müller. Selectivity in Analytical Chemistry. Recommendations for its Use. IUPAC Recommendations 2000
- IUPAC, International Union of Pure and Applied Chemistry (2001) Selectivity in Analytical Chemistry Technical Draft <<http://www.iupac.org/publications/pac/2001/>>

#### UNCERTAINTY

- ISO, Guide to the expression of uncertainty in measurement, International Standards organization, Geneva, 1993
- EURACHEM guide 2nd Edition: Quantifying uncertainty in analytical measurement, <http://www.vtt.fi/ket/eurachem/quam2000-p1.pdf>
- Analytical methods Committee, Analyst, 1995, 120, 2303
- Barwick, V.J., Ellison, S.L.R. (2000) Development and Harmonisation of Measurement Uncertainty Principles. Part (d): Protocol for uncertainty evaluation from validation data, VAM Project 3.2.1, Report n° LGC/VAM/1998/088. LGC, Teddington.

#### IN-HOUSE VALIDATION

- Protocol for in-house method validation, International Union of Pure and Applied Chemistry (IUPAC), 1999
- A.R. Hill, S.L. Reynolds, Analyst, 1999, 124, 953
- IUPAC. International Union of Pure and Applied Chemistry (1995) Nomenclature in Evaluation of Analytical Methods including Detection and Quantification Capabilities, Pure & Appl. Chem., 67, 1699-1723.
- Thompson, M. Wood, R., (1995) Harmonised Guidelines for Internal Quality Control in Analytical Chemistry Laboratories. Pure Appl. Chem., 67, 49-56.

## LÍMITES DETECCIÓN

- L.A. Currie. IUPAC. International Union of Pure and Applied Chemistry (1995) Nomenclature in Evaluation of Analytical Methods including Detection and Quantification Capabilities, Pure & Appl. Chem., 67, 1699-1723.
- C.A. Clayrton et al. Anal. Chem., 55 (1983), 712A-724A.

## VALIDATION OF METHODS

- Harmonised guidelines for in-house validation of methods of analysis (Technical report), International Union of Pure and Applied Chemistry (IUPAC), 1999 <<http://www.iupac.org/.....>>
- Anthony D. Hitchins and Anita Mishra-Szymanski. Qualitative and Quantitative Microbiology Guidelines for Methods Validation . Journal of AOAC INTERNATIONAL, 1999, 82, 402-415 (web: <http://www.aoac.org/pubs/JOURNAL/1999/marapr99.html#STATISTICS>)
- EURACHEM, 'The Fitness for Purpose of Analytical Methods. A Laboratory Guide to Method Validation and Related Topics' EURACHEM Secretariat, Teddington, Middlesex, 1998. (<http://www.vtt.fi/ket/eurachem>)
- AOAC/FAO/IAEA/IUPAC. (2000) Guidelines for Single-Laboratory Validation of Analytical Methods for Trace-Level Concentrations of Organic Chemicals. In Principles and Practices of Method Validation. Edited by A. Fajgelj and A. Ambrus. Royal Society of Chemistry, Cambridge.
- AOAC, (1998) Peer Verified Programs. [http://aoac.org/vmeth/peerverimtd\[1\].htm](http://aoac.org/vmeth/peerverimtd[1].htm)

## EJERCICIOS PRÁCTICOS

### EXACTITUD - TRAZABILIDAD

- 1) Se verificó la trazabilidad de la determinación de cobre en aguas potables por absorción atómica (AAS) utilizando un material de referencia (CRM). El certificado del CRM contenía la siguiente información (ver certificado adjunto):  
CRM (NIST) 1643d de elementos traza en agua  
Concentración de cobre certificada: 20.50 3.80 ppb (ver certificado adjunto)  
Para verificar la trazabilidad, el laboratorio analizó 20 veces el cobre del CRM en condiciones intermedias (en diferentes días y con diferentes analistas). Se obtuvieron los siguientes resultados (en ppb):  
20.57, 16.86, 18.06, 18.47, 17.67, 18.56, 17.77, 17.39, 17.59, 23.74, 24.55, 23.91, 18.78, 21.16, 21.52, 24.17, 20.65, 19.68, 20.73, 18.55  
¿Los resultados proporcionados por el laboratorio son trazables al material de referencia (CRM)?
- 2) Verificar la trazabilidad asociada a la determinación de tetraciclina en músculo utilizando la información generada a partir del análisis de muestras adicionadas. El procedimiento analítico se divide en tres etapas:



La trazabilidad del método se verificó analizando una muestra de músculo a la que se le había adicionado tetraciclina a varios niveles de concentración. La muestra de músculo adicionada no contenía tetraciclina. Se obtuvieron los siguientes resultados:

Tetraciclina adicionada (ppb)	100	200	300	400	500
Resultado (ppb)	67.1	132.3	199.3	275.2	338.5



# Certificate of Analysis

## Standard Reference Material® 1643d

### Trace Elements in Water

This Standard Reference Material (SRM) is intended primarily for use in evaluating methods used in the determination of trace elements in fresh water. SRM 1643d consists of approximately 250 mL of filtered and acidified water in a polyethylene bottle, which is sealed in an aluminized plastic bag to maintain stability. SRM 1643d simulates the elemental composition of fresh water. Nitric acid is present at a concentration of 0.5 mol/L to stabilize the trace elements.

The certified values for 26 elements in SRM 1643d are listed in Table 1. The information values for an additional four elements are provided in Table 2. The analytical methods used for the characterization of this SRM are given in Table 3. All values are reported as mass concentrations [1].

#### NOTICE AND WARNINGS TO USERS

**Expiration of Certification:** This certification of SRM 1643d is valid, within the measurement uncertainty(ies) specified, until **31 July 2003**, provided the SRM is handled in accordance with instructions given in this certificate. This certification is nullified if the SRM is damaged, contaminated, or modified.

**Maintenance of SRM Certification:** NIST will monitor this SRM over the period of its certification. If substantive technical changes occur that affect the certification before the expiration of this certification, NIST will notify the purchaser. Return of the attached registration card will facilitate notification.

**Precautions:** The SRM should be shaken before use because of possible water condensation. To prevent possible contamination of the SRM, do not insert pipets into the bottle. Samples should be decanted at a room temperature of 22 °C ± 5 °C. After use, the bottle should be recapped tightly and returned to the aluminized plastic bag, which should be folded and sealed with sealing tape. This safeguard will protect the SRM from possible environmental contamination and long-term evaporation.

The accuracy of trace element determinations, especially at the µg/L level, is limited by contamination. Apparatus should be scrupulously cleaned and only high purity reagents employed. Sampling and manipulations, such as evaporations, should be done in a clean environment, such as a Class-100 clean hood.

Coordination of the NIST technical measurements was under the direction of J.R. Moody of the NIST Analytical Chemistry Division.

The technical and support aspects involved in the certification and issuance of this SRM were coordinated through the Standard Reference Materials Program by J.S. Kane and B.S. MacDonald.

Willie E. May, Chief  
Analytical Chemistry Division

Thomas E. Gills, Chief  
Standard Reference Materials Program

Gaithersburg, MD 20899  
Revised Certificate Issue Date: 19 March 1999  
See Certificate Revision History on Last Page

Statistical analysis of the experimental data was performed by S.B. Schiller of the NIST Statistical Engineering Division.

The overall coordination of measurements performed by the U.S. Geological Survey, National Water Quality Laboratory, Arvada, CO and by laboratories that participate in the Standard Reference Water Program was under the direction of K. Long.

Table 1. Certified Mass Concentrations for Trace Elements in SRM 1643d

Element	(in mg/L)	Element	(in µg/L)
Calcium	31.04 ± 0.50	Lead	18.15 ± 0.64
Magnesium	7.989 ± 0.035	Lithium	16.50 ± 0.55
Potassium	2.356 ± 0.035	Manganese	37.66 ± 0.83
Sodium	22.07 ± 0.64	Molybdenum	112.9 ± 1.7
Aluminum	127.6 ± 3.5	Nickel	58.1 ± 2.7
Antimony	54.1 ± 1.1	Selenium	11.43 ± 0.17
Arsenic	56.02 ± 0.73	Silver	1.270 ± 0.057
Barium	506.5 ± 8.9	Strontium	294.8 ± 3.4
Beryllium	12.53 ± 0.28	Thallium	7.28 ± 0.25
Boron	144.8 ± 5.2	Vanadium	35.1 ± 1.4
Cadmium	6.47 ± 0.37	Zinc	72.48 ± 0.65
Chromium	18.53 ± 0.20		
Cobalt	25.00 ± 0.59		
Copper	20.5 ± 3.8		
Iron	91.2 ± 3.9		

The mass density, ρ, of SRM 1643d at 22 °C is 1.016 g/mL.

The certified values are equally weighted means of the results of two or more independent analytical methods. The expanded uncertainty [2], whose level of confidence is approximately 95 %, includes within-method sources of uncertainty which were statistically evaluated (Type A) or evaluated by other means (Type B). An allowance for the difference between methods was also included for the following elements: aluminum, barium, beryllium, boron, cadmium, calcium, cobalt, iron, lithium, manganese, nickel, potassium, sodium, strontium, thallium, and vanadium.

Additional elements present in this SRM, other than those certified, are listed below for information only. NIST information values are uncertified values with no uncertainties reported, as there is insufficient information to make an assessment of the uncertainties.

Table 2. Information Mass Concentrations for Selected Trace Elements in SRM 1643d

Element	(in µg/L)	Element	(in mg/L)
Bismuth	13	Silicon	2.7
Rubidium	13		
Tellurium	1		

**PRECISIÓN**

- 1) Para determinar la precisión de la determinación de cobre en aguas por espectroscopia de absorción atómica, se utilizó la información obtenida cuando se verificó la trazabilidad del método utilizando un material de referencia certificado (CRM) de agua potable cuya concentración de cobre certificada era de 20.50 3.80 ppb. Para verificar la trazabilidad del método, el CRM se analizó 20 veces en condiciones intermedias (es decir, los análisis se realizaron en diferentes días, diferentes analistas, etc.). Se obtuvieron los siguientes resultados (en ppb): 20.57, 16.86, 18.06, 18.47, 17.67, 18.56, 17.77, 17.39, 17.59, 23.74, 24.55, 23.91, 18.78, 21.16, 21.52, 24.17, 20.65, 19.68, 20.73, 18.55  
 ¿Cuál es la precisión intermedia del método?

- 2) Una vez que se comprobó la trazabilidad del método de la determinación de cobre en aguas por AAS, se utilizó el método en el análisis de muestras de rutina. Para comprobar que el método estaba bajo control estadístico, se analizaba cada día una muestra control de agua potable. La Tabla 1 muestra los resultados que se obtuvieron después de analizar la muestra control durante 20 días y la Figura 1 muestra el gráfico de control obtenido:

Día	Resultado
1	35.46
2	39.59
3	32.73
4	31.15
5	34.25
6	30.44
7	37.96
8	35.54
9	28.07
10	29.86
11	36.29
12	33.24
13	33.26
14	26.54
15	32.73
16	34.19
17	29.96
18	28.79
19	31.41
20	28.63

Tabla 1. Resultados obtenidos al analizar el cobre de una muestra control de agua potable durante 20 días.

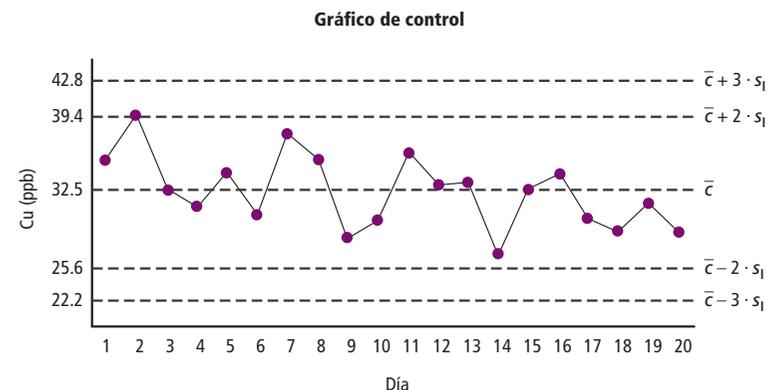


Figura 1. Gráfico de control obtenido al determinar el cobre de una muestra control de agua potable.

**¿Puede obtenerse algún tipo de precisión utilizando la información del gráfico de control? ¿Por qué?**

Calcula el tipo de precisión del método que pueda obtenerse a partir del gráfico de control en el caso de que sea posible.

Posteriormente, se decidió realizar dos replicados cada vez que se analizaba la muestra control. La Tabla 2 muestra los resultados que se obtuvieron después de analizar la muestra control durante 15 días y la Figura 2 muestra el gráfico de control que se obtuvo.

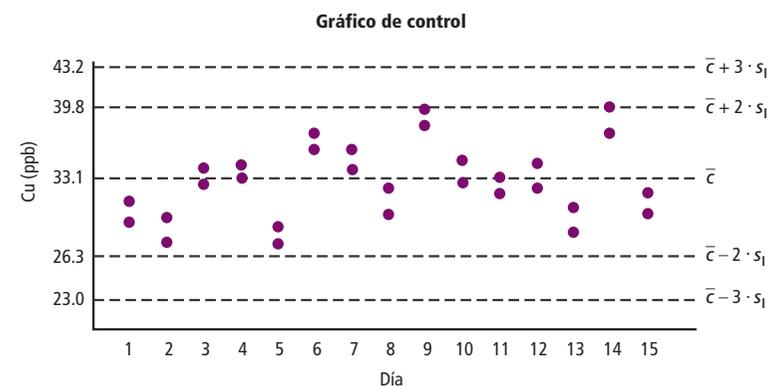


Figura 2. Gráfico de control obtenido al analizar durante 15 días el cobre de la muestra control de agua potable. En cada uno de los días se realizaron dos replicados.

A partir del Análisis de la Varianza (ANOVA) de los resultados de la Tabla 2, calcular la precisión intermedia, la repetibilidad del método y la precisión entre series diferentes. Calcular el límite de repetibilidad del método.

Día	Réplica 1	Réplica 2
1	31.25	29.52
2	27.67	29.85
3	32.86	34.25
4	33.33	34.57
5	29.12	27.56
6	35.95	37.36
7	35.95	34.12
8	32.39	30.12
9	39.45	38.04
10	33.01	34.85
11	31.96	33.46
12	34.60	32.45
13	30.79	28.54
14	39.75	37.38
15	32.10	30.25

Tabla 2. Resultados obtenidos al analizar el cobre de una muestra control de agua potable durante 20 días.

#### INTERVALOS DE LINEABILIDAD

1) Calcular el coeficiente de correlación y realizar el análisis de residuales de la recta de calibración obtenida con los siguientes datos:

x	0,0	1,0	2,0	3,0	4,0	5,0	6,0	7,0	8,0	9,0	10,0
y	0,2	3,6	7,5	11,5	15,0	17,0	20,4	22,7	25,9	27,6	30,2

2) Comprobar, mediante el correspondiente análisis de residuales, que los datos siguientes presentan heteroscedasticidad.

x	0,0	1,0	2,0	3,0	4,0	5,0	6,0
y	0,1	3,8	10,6	13,7	21,6	27,8	28,0

3) Observar la recta de calibrado y los gráficos de residuales al realizar una regresión lineal con los siguientes datos. Compruébese estadísticamente la presencia de algún punto discrepante.

x	0,0	1,0	2,0	3,0	4,0	5,0	6,0
y	0,1	3,8	10,0	14,4	20,7	26,9	29,1

4) La validación de la linealidad de un método cromatográfico para la determinación de un cierto fármaco en sangre se realizó adicionando una cierta concentración de analito a blancos de muestra en el intervalo de 10 a 100 µg/l. El análisis se llevó por duplicado y los resultados fueron los siguientes:

Concentración (µg/l)	Respuesta	
10	13.7	14.8
20	22.2	21.0
30	30.3	32.3
50	51.4	51.6
100	85.9	85.6

Compruébese que el método es lineal en el intervalo estudiado.

#### LÍMITES DE DETECCIÓN Y CUANTIFICACIÓN

1) El plomo en alimentos se analiza mediante GFAAS tras una digestión con microondas de 0.25 g del material y posterior dilución a un volumen final de 25 ml. La recta de calibrado para las disoluciones patrón, entre 1 y 5 ng/ml de Pb, se construyó a partir de las siguientes medidas:

[Pb] (ng/ml)	1,0	2,0	3,0	4,0	5,0
Absorbancia	0,046	0,080	0,102	0,136	0,165

La variabilidad del blanco, incluida la digestión, se determinó con 10 blancos analíticos, obteniéndose los siguientes resultados, expresados en unidades de absorbancia:

1	2	3	4	5	6	7	8	9	10
0,025	0,037	0,012	0,029	0,048	0,026	0,024	0,015	0,041	0,019



**INTI**  
ANIVERSARIO  
1957-2007

Instituto  
Nacional  
de Tecnología  
Industrial



Unión Europea

Proyecto Mejora de la Eficiencia y de la Competitividad de la Economía Argentina

- a) **Calcular el límite de detección del procedimiento:**
- Según el método clásico de la IUPAC (1985)
  - Según las recomendaciones de la IUPAC (1995) con  $\alpha = \beta = 0.05$  y sin tener en cuenta el error asociado a la recta de calibrado.
  - Según las recomendaciones de la IUPAC (1995) con  $\alpha = \beta = 0.05$  y teniendo en cuenta el error asociado a la recta de calibrado.
- b) **Calcular el límite de cuantificación del procedimiento.**

#### SELECTIVIDAD

- 1) La determinación de  $\text{Cu}^{2+}$  en vinos se lleva a cabo mediante espectroscopia de absorción atómica de llama (FAAS). Con objeto de comprobar la ausencia de los efectos de la matriz que afectarían a la exactitud de los resultados, se comparan las rectas de calibrado de patrones acuosos y de adiciones estándar, obteniéndose los siguientes resultados:

PATRONES ACUOSOS		ADICIONES ESTÁNDAR	
[Cu <sup>2+</sup> ] (ppm)	Absorbancia	[Cu <sup>2+</sup> ] (ppm)	Absorbancia
0,0	0,001	0,0	0,020
0,5	0,064	0,5	0,085
1,0	0,136	1,0	0,154
1,5	0,192	1,5	0,220
2,0	0,272	2,0	0,274
2,5	0,332	2,5	0,349

¿Puede concluirse estadísticamente que existe efecto matriz?

# SERVICIO GEOLÓGICO MINERO ARGENTINO (SEGEMAR)

## CÁLCULO DE LA INCERTIDUMBRE EN LOS RESULTADOS ANALÍTICOS



**INTI**

**50**  
ANIVERSARIO  
1957-2007

Instituto  
Nacional  
de Tecnología  
Industrial



Unión Europea

Proyecto Mejora de la Eficiencia y de la Competitividad de la Economía Argentina