Tema 1. Introducción a la Química Analítica

Grupo 2º B. Curso 2025/26

Contenidos

Concepto y relevancia social

El lenguaje de la Química Analítica: Definición y terminología básica

Clasificación de las técnicas y métodos analíticos

Análisis Cualitativo Clásico

Etapas del proceso analítico

Química Analítica Sostenible

Propiedades analíticas

Evaluación y presentación de resultados

1.- Concepto y relevancia social

Dentro de la disciplina científica Química Analítica cabe distinguir entre:

- **Química Analítica:** Ciencia que desarrolla métodos e instrumentos para obtener **información sobre la naturaleza y la composición química** de la materia.
- **Análisis Químico:** Técnica que **aplica los métodos de análisis** desarrollados por la Química Analítica (métodos analíticos) a la resolución de problemas relativos a la naturaleza y composición química de la materia.

Se puede decir que el Análisis Químico está incluido en la Química Analítica, es la parte aplicada. La Química Analítica es una disciplina metrológica puesto que su objeto son los métodos e instrumentos destinados a la medida de información sobre la naturaleza y composición de la materia. La Química Analítica es una ciencia interdisciplinar pues hace uso de conocimientos de Química, principalmente, pero también de Física, Matemáticas, Ingeniería y Biología. Por otro lado, los métodos analíticos son utilizados por otras áreas de la Química y otras ciencias como Bioquímica, Biología o Medicina en sus investigaciones.

La Química Analítica también tiene una gran **relevancia social** en campos muy diversos como:

Industria: Control de calidad de materias primas, procesos y productos.

Comercio: Análisis de mercancías para investigar el cumplimiento de especificaciones legales.

Medioambiente: Evaluación de los niveles de contaminación (aguas, tierras, aire).

Medicina: Análisis clínicos.

Legislación: Establecimiento de normas sobre contenidos de compuestos químicos, vertidos, compuestos tóxicos, etc.

Alimentos: Protección de los consumidores. Establecimiento de alimentos aptos para el consumo.

Análisis forense: Investigación de causas criminales.

2.- **El lenguaje de la Química Analítica**

Terminología básica.

Problema de análisis: Motivo que desencadena el análisis.

Objeto de análisis: Material que interesa analizar.

Muestra: Parte **representativa** de la materia objeto de análisis.

Analito: Especie química (elemento o compuesto) que se investiga (se analiza) en la muestra. Pueden interesar una o más especies químicas, es decir, puede haber uno o más analitos en una muestra.

Interferencias o compuestos interferentes: Especies químicas que acompañan al analito en la muestra y dificultan la obtención de un resultado correcto.

Compuestos inertes: Especies químicas que acompañan al analito en la muestra y que no interfieren en el análisis.

Matriz: Componentes de la muestra exceptuando los analitos. Incluye los compuestos interferentes y los inertes (resto de la muestra).

En relación a **cómo se obtiene la información:**

Análisis: Proceso que proporciona información acerca de la composición o naturaleza de una muestra.

Determinación: Obtención de la identidad (identificación), concentración (cuantificación) o estructura de uno o más analitos, a través del análisis de una muestra.

Técnica analítica: Metodología basada en un principio químico o físico que se utiliza para investigar un analito pues permite obtener una señal característica proporcional a su concentración (volumetría, gravimetría, espectrofotometría, potenciometría, ...).

Método de análisis o método analítico: Serie de operaciones diseñadas para realizar el análisis de una muestra concreta. El método incluye a la técnica o técnicas.

Procedimiento analítico: Conjunto de instrucciones que deben seguirse para aplicar un método de análisis a la determinación de uno o varios analitos en una muestra.

Protocolo: Documento con el conjunto de instrucciones escritas en las que se detalla el procedimiento, y que deben ser seguidas, sin excepción, para que el resultado analítico sea aceptable de forma oficial. Algunos ejemplos son Protocolos de Muestreo, Normas de Análisis, Procedimientos Normalizados de Trabajo y Procedimientos Internos.

Subdivisiones de la Química Analítica

Considerando el **tipo de información**:

- **Análisis Cualitativo:** Identificación (reconocimiento de la identidad) o detección de un analito. Se se investiga la presencia o ausencia de una especie química. El resultado se proporciona en base a la **detección de propiedades químicas y físicas del analito** que permiten deducir si el analito se encuentra en la muestra. La respuesta puede acompañarse de información semicuantitativa dando estimación del posible contenido en analito: "El analito está en concentración superior a..."

- **Análisis Cuantitativo:** Cuantificación del analito. Se investiga la **cantidad o concentración del analito**. Es el análisis más usual.
- **Análisis Estructural:** Investigación de la estructura del analito: la disposición de los átomos en un compuesto químico, los grupos funcionales, y la **distribución espacial y temporal** de los compuestos en la muestra. Suele ser una combinación compleja de Análisis Cualitativo y Cuantitativo.

Considerando la **naturaleza de los analitos**, se distingue entre:

- **Análisis Inorgánico:** Se refiere al análisis de especies químicas inorgánicas (en muestras de suelos, minerales, aleaciones,...), pero también puede realizarse sobre muestras orgánicas, como las muestras biológicas (análisis de heteroátomos o de iones metálicos).
- **Análisis Orgánico:** Se refiere a las especies químicas orgánicas (fármacos, plaguicidas,...). Suele ser más complejo, debido a la mayor cantidad de compuestos y a la similitud de comportamientos. Se ha desarrollado gracias a la aparición de diversas técnicas instrumentales, tales como el Análisis Cromatográfico, la Espectroscopía Infrarroja, la Resonancia Magnética Nuclear y la Espectrometría de Masas.
- **Análisis Bioquímico:** Especies de interés bioquímico (proteínas, enzimas, ADN,...). En los últimos 20 años han ido apareciendo las denominadas ciencias ómicas. El término ómico se utiliza como sufijo para referirse al estudio de una totalidad o conjunto, como genes (Genómica), proteínas (Proteómica) o metabolitos (Metabolómica).

Según **la técnica analítica empleada:**

- **Análisis clásico:** Técnicas basadas en propiedades químicas (Gravimetría, Volumetría, Análisis Cualitativo).
- **Análisis instrumental:** Técnicas basadas en propiedades quimico-físicas (Espectrofotometría, Potenciometría,…).

Según el **campo de aplicación**: Análisis industrial, Análisis clínico, Análisis de alimentos, Análisis forense,…

Niveles de información

La Química Analítica comprende diversas **escalas de trabajo**. Éstas se refieren al analito o a la muestra.

Un **analito** se clasifica en función de su concentración en una muestra en componente:

- **Mayoritario** Si su concentración es mayor del 1% (g/100 g o g/100 mL en disolución acuosa)
- **Minoritario** Si la concentración está entre 1 – 0.01%
  0.01% = 0.01 g/100 g = 100 μg/g o 100 μg/mL = 100 ppm (μg = 10^{-6} g).
- **Traza** (100 μg/mL –1 ng/mL o ppb) (ng = 10^{-9} g)
- **Subtraza** (1 ng/mL – 1 pg/mL o ppt) (pg = 10^{-12} g)
- **Ultratraza** (< 1 pg/mL)

Obsérvese que en estas unidades se utiliza el criterio americano de billón y trillón:

billón = mil millones, trillón = mil billones. Cada una de estas unidades (ppm, ppb y ppt) es mil veces más pequeña que la precedente (factor 10^{-3}): 1 ppm = 10^3 ppb = 10^6 ppt

Por su parte, el **tamaño de la muestra** permite la siguiente clasificación:

- **Macroanálisis** (> 0.1 g)
- **Semimicroanálisis** (0.1 g – 0.01 g)
- **Microanálisis** (0.01 g – 0.001 g)
- **Submicroanálisis** (0.001-0.0001g)
- **Ultramicroanálisis** (< 0.0001 g)

La dificultad en el análisis **aumenta**:

al disminuir la concentración del analito y el tamaño de la muestra,

al aumentar la complejidad de la muestra,

al aumentar la exigencia en la bondad de los resultados.

3.- Clasificación: Técnicas y métodos analíticos

Las técnicas y métodos analíticos se clasifican de acuerdo con distintos criterios.

Por **razones históricas**, se clasifican en:

- **Técnicas y métodos clásicos:** Se trata de los métodos más antiguos, en los que solo se utilizan como instrumentos la balanza y material para medir volúmenes. Basados en propiedades químicas (métodos gravimétricos, volumétricos, análisis cualitativo en tubo de ensayo).
- **Técnicas y métodos instrumentales:** Basados en propiedades químico-físicas (métodos ópticos, eléctricos, ...).
- **Técnicas y métodos de separación:** Permiten la separación del analito de las interferencias para realizar las medidas en condiciones adecuadas. Algunas técnicas permiten realizar también la identificación y cuantificación del analito mediante el acoplamiento con un detector (técnicas cromatográficas y electroforéticas).

De acuerdo con la **propiedad que se mide** (propiedad analítica), las técnicas analíticas se clasifican en:

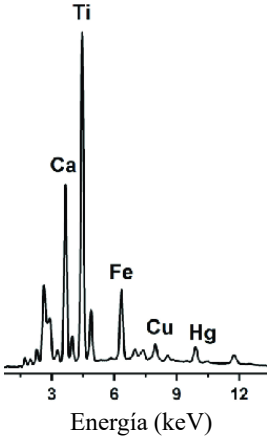
- **Propiedad general:** Volumen: *Volumetría* Peso: *Gravimetría.*
- **Propiedad de la radiación electromagnética:** Absorción de radiación: *Espectrofotometría, absorción atómica*,... Emisión de radiación: *Fluorescencia, misión atómica,...* Dispersión de radiación: *Turbidimetría, nefelometría,...* Resonancia: *Resonancia magnética.*
- **Propiedad eléctrica:** Potencial eléctrico: *Potenciometría.* Corriente eléctrica: *Polarografía, amperometría,...* Resistencia eléctrica: *Conductimetría*
- **Otras propiedades:** Razón carga/masa: *Espectrometría de masas.* Propiedades térmicas: *Métodos térmicos.* Radioactividad: *Activación neutrónica, espectrometría gamma, ...* Velocidad de reacción: *Métodos cinéticos como los métodos enzimáticos.*

4.- Análisis Cualitativo Clásico.

Los métodos cualitativos se basan en la observación de propiedades de las sustancias que permiten deducir su presencia o ausencia en la muestra. Por ello, en el Análisis Cualitativo se tiene finalmente una **respuesta binaria**: SI (presencia del analito en la muestra) o NO (ausencia del analito en la muestra).

Según el tipo de señal que se mide se puede diferenciar entre::

- **Análisis Cualitativo Clásico:** cuando utilizan reacciones químicas para poner de manifiesto la presencia del analito utilizando medidas sensoriales.
- **Análisis Cualitativo Instrumental:** Se utilizan propiedades quimico-físicas que se ponen de manifiesto mediante la medida de señales instrumentales. Las características de la señal como la posición se relacionan con la presencia de un analito concreto. Por ejemplo, en las misiones espaciales los vehículos de exploración van equipados con equipos de análisis elemental como la fluorescencia de rayos X que permite la obtención de señales de energía diferente dependiendo del elemento que emite la fluorescencia. La Espectrometría de Masas es otra técnica de gran valor para identificar la naturaleza de los analitos tanto inorgánicos como orgánicos o incluso de naturaleza bioquímica como proteínas.



En el Análisis Cualitativo Clásico la presencia de analito se determina mediante un **ensayo** de la muestra que se compara con un **ensayo en blanco** (muestra sin analito, indica el NO) y un **ensayo testigo** (muestra con analito que indica el SI). Por ejemplo la identificación de Fe^{3+} puede realizarse añadiendo tiocianato que forma un complejo rojo con el Fe^{3+}. Además de información cualitativa también puede obtenerse información semicuantitativa si se considera la intensidad de la señal. Un problema general es la presencia de **interferentes** que pueden dar señales parecidas y que dificultan la certeza de la identificación.

Características de una reacción para su uso en análisis

- **Cuantitatividad:** Grado de desplazamiento de la reacción hacia la formación de productos:

$$Q\% = \frac{x_{equ}}{x_{max}} \times 100$$

donde Q% es el porcentaje de cuantitatividad de la reacción (0-100), x_{equ} es el grado de avance que ha alcanzado la reacción en el equilibrio y x_{max} el grado de avance máximo que alcanzaría si fuese totalmente cuantitativa, esto es el grado de avance marcado por el reactivo limitante.

Problema 1: Calcular el valor mínimo que debe tener la constante de equilibrio de la reacción $A + R \rightleftarrows P$, para que la cuantitatividad sea igual o mayor del 99.9%, si suponemos que los reactivos están inicialmente a una concentración de 0.1 M.

En primer lugar necesitamos saber el valor del grado de avance máximo. Como ambos reactivos están en proporciones estequiométricas ambos son reactivo limitante y:

$$x_{max} = \frac{\mathrm{C}}{1} = 0.1\ \mathrm{M}$$

Como la cuantitatividad debe ser del 99.9 %, de la formula obtendremos el grado de avance que debe alcanzarse en el equilibrio:

$$x = \frac{Q\%}{100} \times x_{max} = \frac{99.9}{100} \times 0.1 = 0.999 \times \mathrm{C}$$

Por lo tanto, aplicando los balances de grado de avance, las concentraciones en el equilibrio serán:

$$[\mathrm{A}] = \mathrm{C} - x = \mathrm{C} - 0.999 \times \mathrm{C} = \mathrm{C} \times 10^{-3}\ \mathrm{M}$$

$$[\mathrm{R}] = \mathrm{C} - x = \mathrm{C} - 0.999 \times \mathrm{C} = \mathrm{C} \times 10^{-3}$$

$$[\mathrm{P}] = x = 0.999 \times \mathrm{C} \cong \mathrm{C}$$

Sustituyendo se obtiene el valor de la constante de equilibrio:

$$\mathrm{K} = \frac{[\mathrm{P}]}{[\mathrm{A}][\mathrm{R}]} = \frac{\mathrm{C}}{\mathrm{C} \times 10^{-3} \times \mathrm{C} \times 10^{-3}} = \frac{10^6}{\mathrm{C}} = 10^7 \qquad \log \mathrm{K} = 7$$

Por lo tanto, la cuantitatividad depende de C y una reacción como la estudiada con C=0.1 M; debería tener un log K≥7 para ser cuantitativa.

- **Sensibilidad:** Es la concentración mínima que puede identificarse y se conoce como la concentración límite (D suele expresarse μg/mL). Una reacción es muy sensible si D<10 ppm, sensible si D está entre 10 y 100 ppm y poco sensibles si es superior a 100 ppm.
- **Selectividad:** Grado de ausencia de interferencias que presenta la reacción. Según la selectividad la reacción (o el reactivo que la produce) se divide en tres grupos: **Generales** si tiene muchas interferencias porque muchas especies reaccionan con ese reactivo, **selectivas** cuando tiene pocas interferencias (solo un grupo reducido de especies reaccionan con el reactivo) o **específicas** si no tiene interferencias, solo el analito reacciona con el reactivo en las condiciones del ensayo.
- **Seguridad:** Amplitud de condiciones experimentales en las que puede realizarse la reacción. La precipitación del hierro con ferrocianuro es muy segura pues puede realizarse en un amplio intervalo de condiciones y temperatura.

Procedimientos del análisis cualitativo

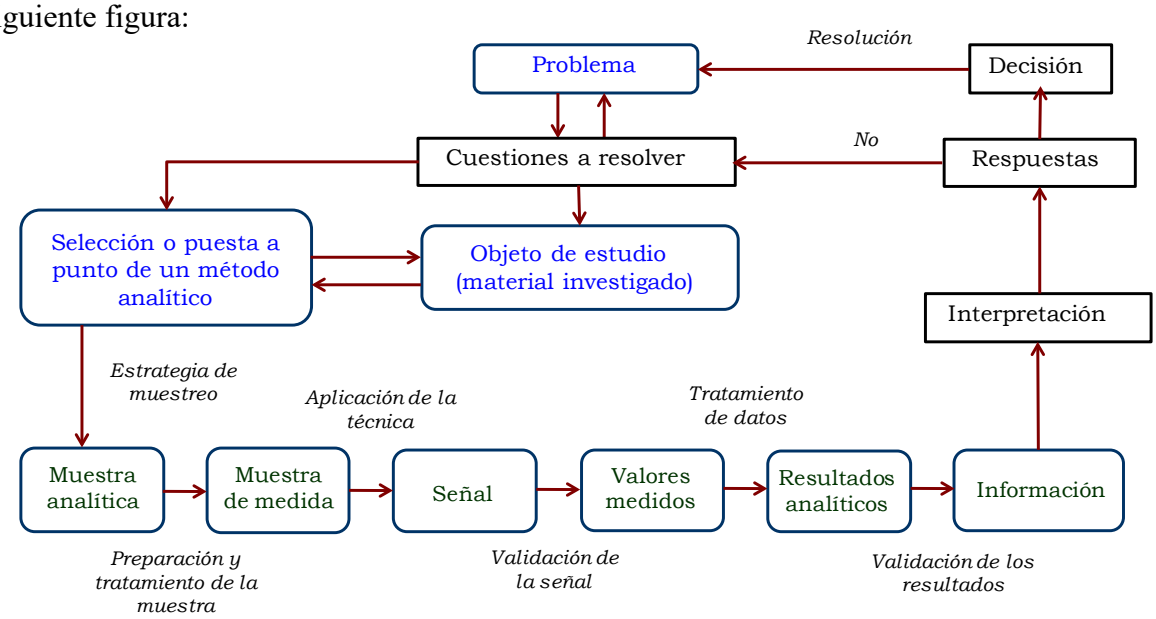
La complejidad de las muestras, por ejemplo un suelo, hace muy probable la presencia de interferencias que reduzcan la certeza de los ensayos de identificación, por ello la principal dificultad dentro del análisis cualitativo, consiste en reconocer una sustancia en presencia de otras de propiedades similares. Se han desarrollado dos procedimientos:

- **Ensayos de identificación directa** en los cuales se eliminan las interferencias específicas de cada analito antes de realizar el ensayo de identificación. Por ejemplo para identificar el Al(III) con naranja de xilenol, debe realizarse una separación previa con NaOH que precipita las interferencias (Fe(III), Cu(II),...) como hidróxidos mientras el aluminio permanece en disolución como aluminato.
- **Marchas** analíticas **sistemáticas**: En este procedimiento se separan las especies en grupos más reducidos mediante reactivos generales de forma que también se reducen las interferencias. En 1841 Fresenius publico la marcha clásica del sulfhídrico basada en la separación por formación de sulfuros y que durante décadas gozó de una gran vigencia.

Su principal inconveniente es la utilización de H_2S que es un gas tóxico y, por ello, se propusieron diversas alternativas para su eliminación entre las que destaca la **marcha del carbonato** desarrollada por Siro Arribas. La marcha del carbonato se inicia con una precipitación con carbonato y separa las especies en seis grupos (Grupo I o soluble en carbonato sódico. Grupo II o insoluble en ácido nítrico. Grupo III o de los cloruros. Grupo IV o de los sulfatos. Grupo V o de los hidróxidos. Grupo VI o de los complejos amoniacales).

5.- Etapas del proceso analítico

El proceso analítico está constituido por el conjunto de operaciones que median entre el objeto a estudiar y el resultado que ha de conducir a la toma de una decisión como se muestra en la siguiente figura:



De forma general, el problema original deriva en un problema analítico cuya resolución implica concretar con claridad cuál es el objeto del análisis, seleccionar el método analítico que se aplicará, proponer una estrategia de muestreo que proporcione una muestra representativa y aplicar el método de análisis con el objetivo de obtener información fiable para resolver el problema planteado. A continuación describiremos de forma breve las distintas etapas:

- **Definición del problema:** Para establecer un procedimiento analítico adecuado deben formularse una serie de cuestiones relativas a: cuestiones a resolver y tipo de análisis requerido, la naturaleza del material investigado (gas, sólido homogéneo, etc.), escala de trabajo, naturaleza del muestreo, naturaleza del analito y de la matriz, importancia de las propiedades analíticas requeridas (métodos de cribado o confirmación) y otras cuestiones (económicas, tiempo, toxicidad, impacto ambiental,...).
- **Selección** del **método de análisis**: Seleccionar el método adecuado es fundamental para el éxito del proceso analítico. Requiere una definición correcta del problema para alcanzar las propiedades analíticas establecidas, atendiendo además a factores complementarios como la disponibilidad de equipo, tiempo, coste, la seguridad y los residuos generados. El método de análisis es el procedimiento para obtener la información, por lo que incluye desde la toma de muestra hasta la presentación de resultados.

- **Toma de muestra**: Se planifica una estrategia de muestreo que proporcione una muestra pequeña pero **representativa** del objeto del análisis y de las cuestiones que debe resolver el proceso analítico. Esta etapa incluye las condiciones de almacenamiento y transporte de la muestra para minimizar la contaminación o las pérdidas de analito.
- **Tratamiento de la muestra:** A continuación, se llevarán a cabo las operaciones de preparación de la muestra que sean necesarias. Con frecuencia es la etapa más larga y requiere un conocimiento de las reacciones químicas y de las propiedades de analitos e interferentes. Pueden reconocerse las siguientes operaciones:

  Preparación: Homogeneización, reducción del tamaño, calcinación, disolución, disgregación o evaporación entre otros procesos.

  Separación: Se separa el analito de las interferencias.

  Conversión o derivatización: Transformación del analito en otra especie química mediante alguna reacción, con frecuencia con un reactivo orgánico.

- **Medida de la señal analítica**: La etapa de medida supone la aplicación de la técnica analítica que proporcionará una o varias señales relacionadas con alguna propiedad del analito o analitos y que, mediante un tratamiento de datos adecuado, deben traducirse en información química sobre la concentración de las especies en estudio. Consta de dos etapas:

  Normalización o calibración: Como corresponde a una ciencia metrológica este paso asegura la fiabilidad de los resultados mediante la obtención de información utilizando patrones (muestras o disoluciones de concentración conocida).

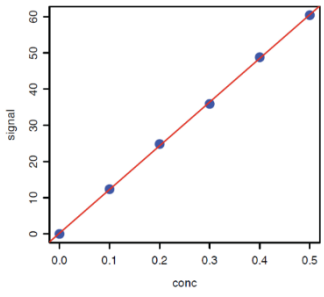
  Determinación: Obtención de información de las muestras de concentración desconocida (las que interesa analizar). La determinación se hace tomando como referencia la información obtenida con los patrones en la etapa anterior.

- **Tratamiento de** los **datos y evaluación de los resultados**: El tratamiento de los datos de calibración permite dar validez a los resultados al relacionar la señal con la concentración de analito. En la mayoría de los métodos se establece una relación lineal:

$$S = m \times C + S_0$$

$$\downarrow$$

$$C_M = \frac{S_M - S_0}{m}$$



- **Validación de los resultados:** Los resultados analíticos deben incluir una estimación de su fiabilidad mediante el cálculo de la incertidumbre de la medida y la valoración en relación con el objetivo del análisis. En laboratorios acreditados deben realizarse además procedimientos de calidad internos (gráficos de control con materiales de referencia, uso de métodos alternativos,...) y externos (auditorías, ejercicios de intercomparación)
- **Presentación de los resultados:** El informe además de comunicar los resultados obtenidos y su incertidumbre con las cifras significativas adecuadas, debe indicar las

limitaciones concretas del método de análisis empleado para asegurar que las conclusiones que se extraigan sean coherentes con los resultados obtenidos.

6. Química Analítica sostenible

La Química Analítica Sostenible es parte de la Química Sostenible y busca la reducción del impacto químico de sus procedimientos sobre el medioambiente y la salud humana. Esto se logra aplicando procesos alternativos y ecológicos basados en 12 principios cuyo objetivo es reducir los riesgos y la huella ecológica de los procedimientos. Estos principios son:

1. Prevenir la generación de residuos.
2. Economía de átomos: lograr procesos que maximicen el uso de los reactivos
3. Síntesis químicas menos tóxicas: utilizar y generar sustancias inocuos o de toxicidad mínima.
4. Diseño de productos químicos seguros: los productos químicos deben tener una toxicidad mínima.
5. Disolventes y auxiliares seguros: Evitar utilizar disolventes o sustancias auxiliares o que sean inocuas.
6. Disminución del consumo de energía.
7. Empleo de materias primas renovables
8. Reducción de productos químicos derivados
9. Uso de procesos catalíticos para acelerar las reacciones
10. Diseño para la degradación: Los productos deben ser degradables en productos inocuos.
11. Análisis de contaminantes en tiempo real
12. Minimización de riesgos de accidentes químicos. Para prevenir accidentes las sustancias utilizadas deben elegirse para minimizar riesgos de accidentes como incendios, explosiones o escapes peligrosos.

7. Propiedades analíticas

Este apartado nos introduce en el tratamiento de datos que corresponde a un área de la Química Analítica denominada Quimiometría. Como se ha estudiado, la Química Analítica es una ciencia metrológica que aplica métodos de análisis principalmente para medir la concentración de uno o varios analitos en una muestra. Para que los resultados obtenidos sean fiables, el método analítico debe estar correctamente validado. La validación consiste en documentar la bondad del método evaluando sus características de calidad o **propiedades analíticas** que permiten conocer la fiabilidad de un método de análisis y compararlo con otros.

Errores en análisis químico

Lo primero que hay que considerar es que toda medida siempre va acompañada de errores lo que reduce la certeza sobre el valor verdadero de la magnitud que se quiere medir. Los errores pueden ser:

- El **error aleatorio** es indeterminado de media cero y supondremos que sigue una distribución normal con desviación estándar σ. Se debe a pequeñas variaciones

incontroladas de variables que afectan al resultado (voltaje, temperatura, densidad,...) y a la incertidumbre asociada a la escala. Varía de forma aleatoria al repetir la medida.

- El **error sistemático** (bías o sesgo) es determinado y no varía al repetir la medida, se debe a variaciones permanentes desconocidas de variables que influyen en el resultado (contaminación de reactivos, temperatura fija diferente a la establecida,...).
- El **error accidental o grueso** es producido por un accidente, equivocación, manipulación incorrecta o mal uso de un instrumento. Da lugar a un valor anómalo que debe ser identificado y eliminado.

Los dos tipos de errores más importantes son el error aleatorio y el error sistemático pues definen la precisión y la exactitud de la medida. Una vez eliminado el error accidental, el resultado de una mediada estará afectado por los dos errores:

$$x_i = \mu + \delta + \varepsilon_i$$

donde x_i es el resultado del análisis i de la misma muestra, μ el valor verdadero, δ el error sistemático y ε_i el error aleatorio. Si realizamos un número grande de medidas y tomamos la media:

$$\bar{x} = \frac{\sum x_i}{N} = \mu + \delta + \frac{\sum \varepsilon_i}{N} = \mu + \delta \qquad \rightarrow \qquad \delta = \bar{x} - \mu$$

pues la media del error aleatorio es cero. Así, un resultado individual puede ponerse como:

$$x_i = \mu + (\bar{x} - \mu) + (x_i - \bar{x})$$

el error sistemático se obtiene de la diferencia entre la media de un número grande de medidas y el valor verdadero, mientras que el error aleatorio se relaciona con la diferencia entre la medida y la media y se medirá mediante la desviación estándar. Estas relaciones permitirán evaluar algunas propiedades analíticas, pero primero debemos estudiar la medida de la incertidumbre que genera el error aleatorio.

Medida de la incertidumbre

La existencia de errores aleatorios introduce incertidumbre en el resultado, de forma que medidas repetidas no dan el mismo valor al verse afectadas por el error aleatorio:

$$x_i = \mu + \varepsilon_i$$

Por ello, un resultado afectado de un error aleatorio es una variable aleatoria que se distribuirá según una distribución de probabilidad. La estadística se basa en la incertidumbre que introducen los errores aleatorios lo que hace que el igual se ensanche a un intervalo que depende de la magnitud del error aleatorio. Así, si conocemos un valor exacto μ=1.50, podemos decir que μ≠1.55, pero si es medimos una variable aleatoria, podemos obtener 1.49 una vez y 1.52 la siguiente por lo que $x = 1.49 = 1.52$ pues ambos valores son la medida de la misma realidad (μ).

Para detectar y cuantificar el error aleatorio es necesario realizar medidas repetidas y establecer una estimación del valor central (media) y una estimación de la dispersión (desviación estándar). Se llama **población** al número infinito de todos las posibles medidas que puedan obtenerse. La población se caracteriza por una **distribución de probabilidad** que asigna a cada

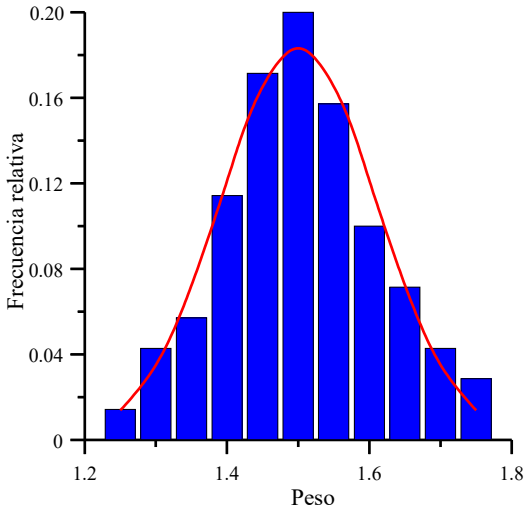
valor la probabilidad de que ocurra. La distribución de probabilidad más usual de los errores aleatorios es la **distribución normal** o gaussiana que tiene la forma:

$$f(x_i) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2}$$

donde el error aleatorio de cada medida es:

$$\varepsilon_i = x_i - \mu$$

Esta distribución está normalizada (la probabilidad total, área bajo la curva, es uno), es decir, la probabilidad de obtener cualquier valor es 1 o del 100%. La gaussiana está definida por dos parámetros, la **media de la población** o valor central μ indica el valor de máxima probabilidad, y la desviación estándar σ, que mide la dispersión de los valores y está relacionada con la anchura de la distribución (para la distribución normal la anchura al 60% de altura vale 2 σ). Los valores del valor central y de la desviación estándar definen la población de todos los valores que podemos obtener. La **muestra** es el conjunto finito de medidas realizadas y a partir de la cual se extraen conclusiones acerca de la población. Por ejemplo, supongamos que pesamos una pesa de 1.50 g y que la población de medidas sigue una distribución N(1.5, 0.11), si obtenemos una muestra de 50 pesadas veremos resultados diferentes y si representamos la frecuencia relativa con la que ha aparecido cada valor obtendremos un **histograma** que es la representación de la población que se realiza con la muestra. En la figura se muestra el histograma obtenido con la muestra (barras azules) y la distribución gaussiana de la población, (línea roja).



Los parámetros de la población μ y σ no se conocen y para realizar una estimación debemos obtener una muestra de N medidas. La media de los valores de la muestra nos dará una estimación del valor central y la desviación típica una estimación de la de desviación estándar que mide la magnitud del error aleatorio:

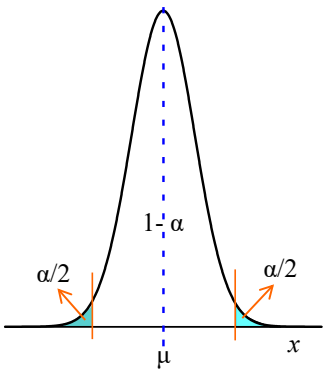
$$s = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N-1}}$$

Cuanto mayor es *s*, mayor es la dispersión de los resultados medidos. La desviación estándar también puede cuantificarse de forma relativa respecto a la media calculando la desviación estándar relativa (DER) también conocido como coeficiente de variación (CV):

$$\text{CV\%} = \frac{s}{\bar{x}} \times 100$$

Como el área bajo la distribución está relacionada con la probabilidad de obtener unos valores de la variable aleatoria, es posible establecer un intervalos alrededor de la media verdadera y relacionarlo con la probabilidad de obtener los valores dentro del mismo:

$$x \equiv \mu \pm z_{\alpha/2} \times \sigma$$



donde μ y σ son la media y la desviación estándar de la población, *z* el estadístico de la normal tipificada y α el **nivel de significación**, que indica la probabilidad de los valores que quedan fuera del intervalo, de forma que (1-α)×100 es el nivel de confianza o probabilidad en porcentaje de obtener un valor que esté en el intervalo. Por ejemplo, la probabilidad de encontrar un valor entre μ-σ y μ+σ es del 68.3% (z=1) y entre μ-2σ y μ+2σ, del 95.5% (z=2). Como la distribución normal es simétrica alrededor del valor verdadero, fuera del intervalo quedan dos colas de probabilidad α/2. Por ejemplo, si α=0.05, de las tablas obtenemos $z_{0.025}$=1.96 y el 95% de las medidas se situarán en el intervalo:

$$x \equiv \mu \pm 1.96 \times \sigma$$

y quedaran fuere el 5%, 2.5% por debajo y 2.5% por encima.

Tabla de valores de z para diferentes niveles de significación de una cola.

α	0.5	0.4	0.3	0.2	0.1	0.05	0.025	0.01
z_α	0.000	0.253	0.524	0.842	1.281	1.645	1.960	2.325

También podemos utilizar este intervalo para construir un intervalo de confianza que mida la incertidumbre de la medida realizada. Si conocemos la desviación estándar de la población:

$$\mu \equiv x \pm z_{\alpha/2} \times \sigma$$

para una medida individual y

$$\mu \equiv \bar{x} \pm z_{\alpha/2} \times \sigma_{\bar{x}} = \mu \equiv \bar{x} \pm z_{\alpha/2} \times \frac{\sigma}{\sqrt{N}}$$

para una media de *N* medidas.

Cuando la desviación estándar verdadera de la población no se conoce (no ha sido estimada con más de 30 medidas), debe aplicarse el estadístico *t* de Student:

$$\mu \equiv \bar{x} \pm t_{\alpha/2,N-1} \times \frac{s}{\sqrt{N}}$$

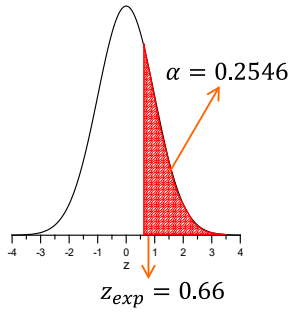
donde *s* es la desviación estándar muestral estimada con las *N* medidas. El intervalo de confianza indica la incertidumbre de la medida pues cualquier valor dentro del mismo podría ser el valor verdadero con una probabilidad definida por el nivel de significación. Por ejemplo si α=0.05, puede considerarse que existirá una probabilidad del 95 % de que el valor verdadero esté dentro del intervalo de confianza y una probabilidad del 5% de que se encuentre fuera del mismo. En realidad esto es cierto solo si se realizan un número muy elevado de intervalos de confianza, en el 95% de ellos el valor verdadero estará dentro.

Problema 2. Se realizan 80 pesadas de una pesa y se calcula μ=10.000 g y σ=1.51 mg. Una pesada individual da 10.001 g. ¿Cuántas medidas estarán por encima de este valor?

Tendremos que obtener el valor z del valor frontera:

$$z_i = \frac{x_i - \mu}{\sigma} = \frac{10.001 - 10.000}{1.5 \times 10^{-3}} = 0.667$$

En la tabla de z se comprueba que este valor corresponde con una probabilidad de 0.255, luego el 25.5% de las medidas quedaran por encima del valor 10.001 g



Propiedades analíticas

Las propiedades analíticas se clasifican en tres grupos jerarquizados: propiedades supremas, básicas y complementarias.

Las **propiedades supremas**, son la base de la calidad de los resultados analíticos:

- **Representatividad**: Grado de concordancia de la muestra con el objeto del análisis y con el problema analítico.
- **Exactitud**: Grado de concordancia entre el resultado y el valor verdadero. De conocerse el valor verdadero, se podría medir como error absoluto, relativo o como recuperación:

$$\mathrm{E} = x_i - \mu \qquad \mathrm{E_r} = \frac{x_i - \mu}{\mu} \times 100 \qquad \mathrm{R(\%)} = \frac{x_i}{\mu} \times 100 = \mathrm{E_r} + 100$$

  donde x_i es el resultado del análisis, que puede ser una medida individual o la media de un número pequeño de medidas.

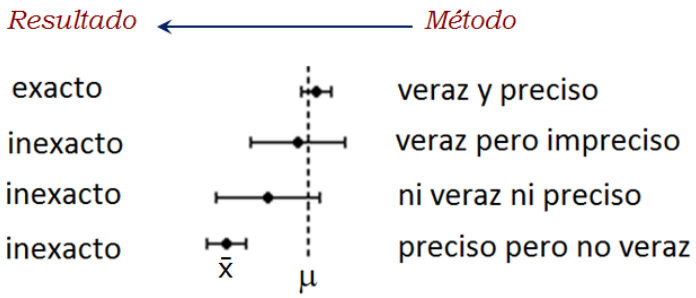
Un resultado debe cumplir ambas propiedades para ser aceptable, pues un resultado exacto pero poco representativo es de poca calidad, al igual que uno representativo pero inexacto. Es importante indicar que estas propiedades en general no pueden medirse pues se desconoce el valor verdadero, razón por la que se realiza el análisis. Para asegurarlas debe aplicarse un método correctamente validado que cumpla las siguientes **propiedades básicas**, que son la base de la calidad del método de análisis:

- **Muestreo adecuado**: Indica que el muestreo se ha realizado según unas normas estadísticas que aseguran su coherencia con el objeto del análisis. No puede asegurarse la representatividad sin muestreo adecuado.
- **Veracidad**: Grado de concordancia entre el valor medio de una serie grande de resultados y el valor de referencia aceptado como verdadero. Cuantifica el error sistemático:

$$\mathrm{E} = \bar{x} - \mu \qquad \mathrm{E_r} = \frac{\bar{x} - \mu}{\mu} \times 100 \qquad \mathrm{R(\%)} = \frac{\bar{x}}{\mu} \times 100 = \mathrm{E_r} + 100$$

- **Precisión**: Grado de concordancia entre los datos obtenidos de una serie de resultados. Refleja el efecto de los errores aleatorios producidos durante el proceso analítico. Se cuantifica mediante la desviación estándar. Cuanto mayor es la desviación estándar, menor es la precisión y mayor la incertidumbre.

  No puede asegurarse la exactitud de un resultado si el método no es veraz y preciso:



Un método veraz pero impreciso podría dar resultados inexactos pues no puede saberse en qué zona del intervalo de confianza estará el resultado.

Con fines prácticos en el estudio de métodos analíticos, dentro de la precisión puede distinguirse entre:

Repetibilidad: En la precisión de un método en un laboratorio. Se mide como la cercanía entre sí de las medidas obtenidas con el mismo método, sobre la misma muestra, en las mismas condiciones (operador, laboratorio, instrumental,…) y en un intervalo pequeño de tiempo. Se hacen de 6 a 8 réplicas, se calcula la desviación estándar. Si el blanco es una fuente de variación también se replica. Debe realizarse a varios niveles de concentración.

Reproducibilidad: Es la precisión de un método aplicado en varios laboratorios. Se mide como la cercanía entre sí de las medidas obtenidas con el mismo método, sobre la misma muestra, en laboratorios diferentes por operadores diferentes utilizando equipos diferentes. Se mide en estudios de intercomparación de laboratorios:

$$s_R^2 = s_L^2 + s_r^2$$

Como cada laboratorio puede tener un error sistemático diferente, estos errores contribuyen como componentes aleatorios entre todos los laboratorios. Así, la varianza de reproducibilidad (s_R^2) es mayor que la de repetibilidad (s_r^2) al incorporar la varianza interlaboratorio (s_L^2).

- **Sensibilidad**: Capacidad para discriminar entre pequeñas diferencias de concentración del analito. La sensibilidad se cuantifica mediante la pendiente de la recta de calibrado a la concentración de interés.

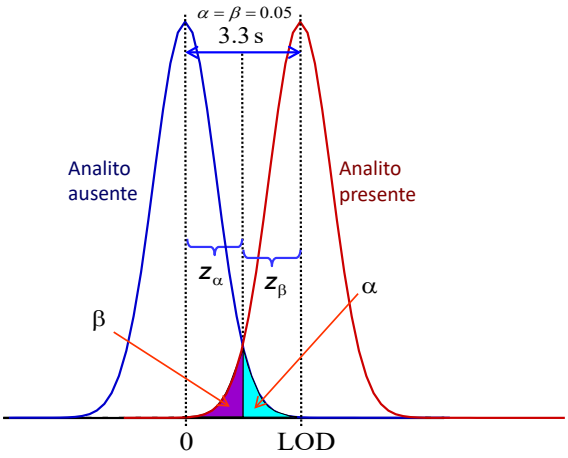
$$S = m \times C + S_0 \qquad \text{Sensibilidad} = \frac{\Delta S}{\Delta C} = m$$

- **Límite de detección** (LOD): Es la concentración mínima que proporciona una señal significativamente distinta del blanco con un nivel de significación dado:

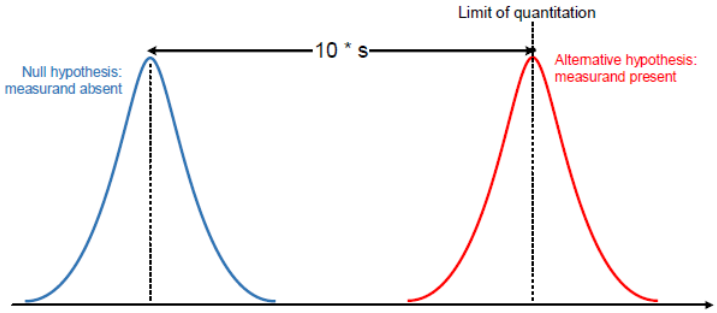
$$S_{LOD} = S_0 + k \times \sigma_0 = m \times C_{LOD} + S_0 \qquad \rightarrow \qquad LOD \equiv C_{LOD} = \frac{k \times \sigma_0}{m}$$

En 1978 se propuso que k = 3 (criterio 3s). En la actualidad se realiza un tratamiento estadístico más riguroso que depende de los niveles de significación.

Por ejemplo, si k = 1.64 + 1.64 = 3.3 se tiene el límite de detección que indica la mínima concentración de analito que es claramente distinta del blanco con α = 5% y β = 5%.



- **Límite de cuantificación** (LOQ): Concentración mínima que se puede cuantificar (se considera el límite inferior del rango lineal). Similar al LOD pero con un valor de k = 10.



- **Selectividad**: Cuantifica el grado de ausencia de interferencias. Las interferencias son debidas a otras especies contenidas en la muestra.

Las propiedades analíticas supremas, exactitud y representatividad, son la base de la calidad de los resultados del método de análisis y se fundamentan en las propiedades básicas. Así, la exactitud está condicionada por la precisión, que mide la magnitud de los errores aleatorios, y la veracidad que mide la magnitud de los errores sistemáticos, mientras que la representatividad está condicionada por el muestreo adecuado. Sin conseguir las propiedades básicas no pueden alcanzarse las supremas. El proceso analítico también se ve afectado por otras propiedades aparentemente menos significativas, que se denominan complementarias como el coste, la seguridad o el impacto medioambiental. Sin embargo, estas propiedades pueden tener en ocasiones un papel fundamental.

Problema 3. En la validación de un método de determinación del insecticida permetrina se analiza un MRC (material de referencia certificado) cuyo valor certificado es 3.45 ± 0.02 µg/Kg. La aplicación del método diez veces dio los siguientes resultados: 3.38; 3.23; 3.29; 3.27; 3.21; 3.35; 3.28; 3.30; 3.25 y 3.40 µg/Kg. Calcular y discutir los parámetros de precisión, veracidad e incertidumbre. Estudia la exactitud del último resultado.

Calculamos la media y la desviación estándar de la muestra:

$$\bar{x} = \frac{1}{\mathrm{N}}\sum_{i=1}^{\mathrm{N}} x_i = 3.296 \qquad s = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N-1}} = 0.063$$

La precisión se evalúa por la desviación estándar y el CV%:

$$\mathrm{CV\%} = \frac{s}{\bar{x}} \times 100 = \frac{0.063}{3.296} \times 100 = 1.9\ \%$$

El método tiene una precisión aceptable ($< 2\%$). La incertidumbre de la media se obtiene con el intervalo de confianza:

$$\mu \equiv \bar{x} \pm t_{\alpha/2,N-1} \times \frac{s}{\sqrt{N}}$$

Como $t_{0.025,9} = 2.262$: $\quad \mu \equiv 3.251 - 3.341$

La veracidad:

$$E = \bar{x} - \mu = 3.296 - 3.45 = -0.154 \qquad E_r\% = \frac{\bar{x} - \mu}{\mu} \times 100 = -4.67\%$$

La veracidad es baja, podría existir error sistemático (esto se confirma porque el valor verdadero no está dentro del intervalo de confianza).

La exactitud del último resultado:

$$E = x_i - \mu = 3.40 - 3.45 = -0.05 \qquad E_r\% = \frac{x_i - \mu}{\mu} \times 100 = -1.45\%$$

Es el dato más exacto, el menos exacto es:

$$E = x_i - \mu = 3.21 - 3.45 = -0.24 \qquad E_r\% = \frac{x_i - \mu}{\mu} \times 100 = -6.96\%$$

8. Evaluación y presentación de resultados

Además de evaluar la calidad de los resultados y del método mediante la determinación de las propiedades analíticas, los resultados de los análisis permiten tomar decisiones estadísticas en relación con la finalidad del análisis. Por ejemplo, puede realizarse una determinación de nitrato en agua para saber si cumple la normativa y es potable. La inferencia estadística consiste en una serie de contrastes o ensayos de hipótesis utilizando pruebas estadísticas que permiten tomar decisiones sobre la población. Estas pruebas realizan análisis comparativos, por ejemplo comparar la media con un valor dado o dos medias entre sí. Una hipótesis estadística es una afirmación sobre los parámetros de una o más poblaciones:

Hipótesis nula, H_0: Afirmación que se quiere contrastar y es la que se acepta o rechaza.

Hipótesis alternativa, H_1: Es la hipótesis alternativa a H_0, si se acepta H_0, se rechaza H_1 y viceversa.

El procedimiento de prueba de una hipótesis consiste en utilizar la información de una muestra aleatoria obtenida de la población de interés. Si la información es consistente con H_0, se concluye que la hipótesis es cierta y se acepta. Si la información es inconsistente se concluye que la hipótesis es falsa y se rechaza.

Errores que se pueden cometer en un ensayo estadístico

Error tipo I: Error que se comete al rechazar H_0 siendo cierta. Su probabilidad es el nivel de significación (α).

Error tipo II: Error que se comete cuando se acepta H_0 siendo falsa, su probabilidad se denota por β.

Nivel de significación (α): Probabilidad de que al rechazar la H_0 esta sea correcta.

Es muy importante entender el nivel de significación (α): probabilidad de equivocarse al rechazar H_0 y nivel de error tipo I que estamos dispuestos a aceptar. Es necesario establecer el valor de α pues las colas de las distribuciones llegan al infinito y pretender una certeza absoluta nos llevaría a tener que aceptar cualquier valor.

Dependiendo de cómo se asigne esta probabilidad α de acuerdo con H_0, el ensayo puede ser de una cola o de dos

Test de dos colas H_0: $\mu = \mu_o$ H_1: $\mu \neq \mu_o$. Se consideran erróneos resultados alejados de la media tanto por encima y como por debajo por lo que la probabilidad α se coloca en las dos colas de la distribución, α/2 en cada una.

Test de una cola H_0: $\mu \leq \mu_o$ H_1: $\mu > \mu_o$. Toda la probabilidad α se coloca en una cola de la distribución pues solo son erróneos los que están por encima o los que están por debajo según sea H_0.

Los estudios de comparación estadística se basan en calcular un estadístico con los datos experimentales y compararlo con el obtenido en la tabla correspondiente para un nivel de significación dado. Pasos en un estudio comparativo:

1. Especificar la hipótesis nula y la alternativa
2. Especificar el estadístico apropiado a partir de H0
3. Definir el nivel de significación (esto es, definir que consideramos un resultado excepcional)
4. Calcular el estadístico y compararlo con los valores críticos tabulados

La **hipótesis estadística** se relaciona con una variable aleatoria y su desviación estándar. Por ejemplo en una prueba de dos colas se determina si el valor que se quiere comprobar está dentro del intervalo de confianza de la variable (entonces se acepta H_0: $\mu = \mu_o$) o fuera (entonces se rechaza H_0 y se acepta H1: $\mu \neq \mu_o$). Por ello, para realizar la prueba de hipótesis necesitamos un valor de la varianza de la variable que se quiere estudiar. En ocasiones esta variable es función de otras y para obtener su varianza se hace uso de la ecuación de **propagación de errores**. Cuando las varianzas son independientes la ecuación es:

$$f = f(x, y, z, \ldots) \quad \sigma_f^2 = \left(\frac{\partial f}{\partial x}\right)^2 \sigma_x^2 + \left(\frac{\partial f}{\partial y}\right)^2 \sigma_y^2 + \left(\frac{\partial f}{\partial z}\right)^2 \sigma_z^2 + \cdots$$

Problema 4. Obtén la varianza de la media de N valores independientes de la variable *x*.

La media es:

$$\bar{x} = \frac{1}{\mathrm{N}} \sum_{i=1}^{\mathrm{N}} x_i$$

Por lo tanto:

$$\sigma_{\bar{x}}^2 = \sum_{i=1}^{\mathrm{N}} \left(\frac{\partial \bar{x}}{\partial x}\right)^2 \sigma_x^2 = \sum_{i=1}^{\mathrm{N}} \left(\frac{1}{\mathrm{N}}\right)^2 \sigma_x^2 = \frac{\sigma_x^2}{\mathrm{N}}$$

Se ha tenido en cuenta que $\sum_{i=1}^{\mathrm{N}} 1 = \mathrm{N}$ y que σ_x^2 es igual para todos los valores pues pertenecen a la misma población.

***Rechazo de resultados anómalos**:* Un dato anómalo es aquel cuyo valor es muy superior o inferior al de los datos restantes debido a un error grueso. Aplicaremos la **prueba de Dixon** que asume la normalidad de la distribución y compara la distancia entre los datos y el cuestionado. El dato cuestionado (x_Q) que será el más pequeño o el más grande, se compara con el dato más próximo (x_P) y el más lejano (x_L). Se calcula el estadístico:

$$Q_o = \frac{|x_Q - x_P|}{|x_Q - x_L|}$$

Este valor se compara con el de la tabla y si $Q_o > Q_{1-\alpha,N}$ se rechaza el dato cuestionado.

N	3	4	5	6	7	8	9	10
95%	0.970	0.829	0.710	0.625	0.568	0.526	0.493	0.466
99%	0.994	0.926	0.821	0.740	0.698	0.634	0.598	0.568

Otra forma es aplicar la **prueba de Grubbs** en la que se compara el dato cuestionado con la media y la desviación estándar de todos los datos incluido el cuestionado. El estadístico es:

$$G_o = \frac{|x_Q - \bar{x}|}{s}$$

Este valor se compara con el de la tabla y si $G_o > G_{1-\alpha,N}$ se rechaza el dato cuestionado.

N	3	4	5	6	7	8	9	10
95%	1.154	1.481	1.715	1.887	2.020	2.127	2.215	2.290
99%	1.155	1.496	1.764	1.973	2.139	2.274	2.387	2.482

Problema 5: En la validación de un nuevo método de análisis se realizaron 6 análisis de una muestra certificada obteniendo los siguientes resultados 5.20, 5.12, 5.29, 5.22, 5.03 y 5.25 en mg/L. Estudia si se debe rechazar algún resultado y calcula el valor que sería rechazable.

El dato cuestionado debe ser el más pequeño 5.03 o el más grande 5.29. Probaremos el primero pues está más alejado. Si aplicamos Dixon:

$$Q_o = \frac{5.12 - 5.03}{5.29 - 5.03} = \frac{0.09}{0.26} = 0.346 < Q_{0.95,6} = 0.625 \quad No\ se\ rechaza$$

El valor del dato que debería rechazarse sería

$$\frac{5.12 - x_Q}{5.29 - x_Q} = Q_{0.95,6} = 0.625 \qquad x_Q = 4.84$$

Para aplicar Grubbs, necesitamos la media y la desviación estándar de los datos:

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i = 5.185 \qquad s = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N-1}} = 0.0948$$

Por lo tanto:

$$G_o = \frac{5.185 - 5.03}{0.0948} = 1.635 < G_{0.95,6} = 1.887 \quad No\ se\ rechaza$$

En este caso obtener el primer dato que se rechazaría es más complicado pues afecta a la media y a la desviación estándar, pero mediante un proceso iterativo se obtiene $x_Q = 4.88$

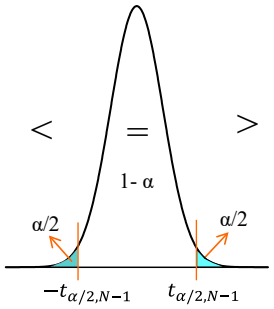
Comparación de un resultado con un valor de referencia: Esta prueba permite establecer si existe error sistemático en un método o si una muestra cumple un límite establecido. A partir del intervalo de confianza se establece el estadístico, por ejemplo para la media:

$$\mu_0 = \bar{x} \pm t_{\alpha/2,N-1} \times \frac{s}{\sqrt{N}} \qquad \rightarrow \qquad \mathrm{T_o} = \frac{\bar{x} - \mu_0}{s/\sqrt{N}}$$

La prueba puede ser:

- De **dos colas** cuando no importa si es diferente por defecto o por exceso. Por ejemplo en la detección del error sistemático:

$$\mathrm{H_o}: \mu = \mu_o \quad \mathrm{H_1}: \mu \neq \mu_o$$



En este caso el nivel de significación elegido se reparte entre las dos colas de la distribución t de Student y el estadístico $\mathrm{T_o}$ se compara con los valores límites de la distribución:

$$si \quad -t_{\alpha/2,N-1} < \mathrm{T_o} < t_{\alpha/2,N-1} \quad se\ acepta\ H_o$$

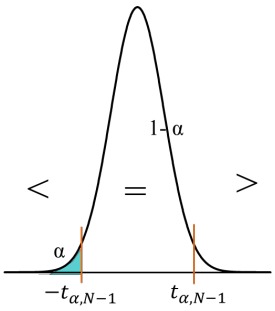
puesto que la distribución t de Student es simétrica. En este caso también puede evaluarse el valor absoluto:

$$si \quad |T_o| < t_{\alpha/2,N-1} \quad se\ acepta\ H_o$$

Pues tanto la distribución z como la t son simétricas alrededor de cero.

- De **una cola** cuando importa que el valor sea mayor o menor. Por ejemplo en el rechazo de una muestra que supera el límite establecido por la ley de una sustancia

$$\mathrm{H_o}: \mu < \mu_o \quad \mathrm{H_1}: \mu \geq \mu_o$$
$$si\ \mathrm{T_o} < -t_{\alpha,N-1}\ se\ acepta\ \mathrm{H_o}$$

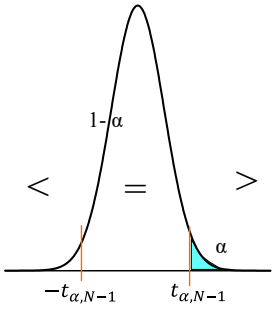


Pero pueden encontrarse otras posibilidades y dependiendo de ellas se debe elegir el estadístico tabulado en una cola o en la otra de la distribución:

$$\mathrm{H_o}: \mu \leq \mu_o \quad \mathrm{H_1}: \mu > \mu_o$$
$$si\ \mathrm{T_o} < +t_{\alpha,N-1}\ se\ acepta\ \mathrm{H_o}$$

$$\mathrm{H_o}: \mu > \mu_o \quad \mathrm{H_1}: \mu \leq \mu_o$$
$$si\ \mathrm{T_o} > +t_{\alpha,N-1}\ se\ acepta\ \mathrm{H_o}$$

$$\mathrm{H_o}: \mu \geq \mu_o \quad \mathrm{H_1}: \mu < \mu_o$$
$$si\ \mathrm{T_o} > -t_{\alpha,N-1}\ se\ acepta\ \mathrm{H_o}$$



Este procedimiento de comparación es similar cuando la variable aleatoria sigue otras distribuciones, por ejemplo la normal utilizando el estadístico *z* o la Chi2 para comparar la varianza con un valor utilizando el estadístico χ^2., en este caso la distribución no es simétrica y deben encontrarse en las tablas los dos valores del estadístico de ambas colas. Además, el

estadístico *z* puede utilizarse en comparaciones de variables que siguen otras distribuciones si se aproximan a la normal como la t de Student si N>30 o la de Poisson si N es suficientemente grande.

Tabla de valores de t de Student para diferentes grados de libertad y tres niveles de significación

N-1	2	3	4	5	6	7	8	9	z
0.05	2.920	2.353	2.132	2.015	1.943	1.895	1.860	1.833	1.645
0.025	4.303	3.182	2.776	2.571	2.447	2.365	2.306	2.262	1.960
0.01	6.965	4.541	3.747	3.365	3.143	2.998	2.896	2.821	2.326

Problema 6: Se sospecha que un método de valoración ácido-base tiene un error de indicador significativo y, por tanto, tiende a dar resultados con error sistemático. Para comprobar esto se valora una muestra de concentración conocida 0.0524 M. Contrasta la hipótesis de falta de exactitud del método si se obtuvieron los siguientes resultados: 0.0519, 0.0520, 0.0518, 0.0521, 0.0519, 0.0515

No hay ningún dato anómalo por lo que la media y la desviación estándar es:

$$\bar{x} = 0.05187 \qquad s = 0.000207$$

Como se quiere estudiar si hay sesgo positivo o negativo, el test es de dos colas y H_o: $\mu = \mu_o$:

$$T_o = \frac{|\bar{x} - \mu_0|}{s} \times \sqrt{N} = \frac{|0.05187 - 0.0524|}{0.000207} \times \sqrt{6} = 6.272$$

$$\text{Como } T_o = 6.272 > t_{0.025,5} = 2.571 \quad Si\ hay\ sesgo$$

Comparación de dos resultados: Para comparar dos resultados se aplica el mismo procedimiento pero con la variable diferencia. Si nos centramos en la comparación de dos medias:

$$d = \bar{x}_2 - \bar{x}_1 \quad \rightarrow \quad s_d^2 = s_{\bar{x}_1}^2 + s_{\bar{x}_2}^2 = \frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}$$

Y el estadístico para comparar las medias ($d = 0$) será:

$$T_o = \frac{\bar{x}_1 - \bar{x}_2}{s_d}$$

Sin embargo existe el problema de los grados de libertad para buscar el estadístico en las tablas. Los grados de libertad dependen de si las varianzas de ambas medias son homogéneas o no. Si son homogéneas provienen de la misma población de errores aleatorios y miden el mismo valor. Cuando esto sucede, puede obtenerse una estimación más exacta de la varianza utilizando los errores aleatorios de ambas medias:

$$s^2 = \frac{(N_1 - 1)s_1^2 - (N_2 - 1)s_2^2}{N_1 + N_2 - 2}$$

Ahora los grados de libertad son $v = N_1 + N_2 - 2$ y la varianza de la diferencia:

$$s_d^2 = \frac{s^2}{N_1} + \frac{s^2}{N_2} = s^2\left(\frac{1}{N_1} + \frac{1}{N_2}\right)$$

Sustituyendo en el estadístico:

$$\mathrm{T_o} = \frac{\bar{x}_1 - \bar{x}_2}{s \times \sqrt{\frac{1}{N_1} + \frac{1}{N_2}}}$$

Las hipótesis pueden ser iguales que en el caso anterior, por ejemplo:

$$\mathrm{H_o}: \mu_2 - \mu_1 = 0 \quad \mathrm{H_1}: \mu_2 \neq \mu_1$$

$$Si \quad -t_{\alpha/2,\, N_1+N_2-2} < \mathrm{T_o} < t_{\alpha/2,\, N_1+N_2-2} \qquad se\ acepta\ H_o$$

Comparación de varianzas: Para conocer si dos varianzas son homogéneas y por lo tanto provienen de la misma población, se recurre a un test F. La distribución F es la que sigue el cociente de dos varianzas. El estadístico es:

$$\mathrm{F_o} = \frac{s_1^2}{s_2^2}$$

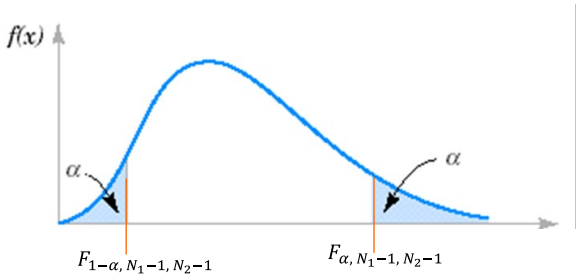
Y la hipótesis nula:

$$\mathrm{H_o}: \sigma_1^2 = \sigma_2^2 \quad \mathrm{H_1}: \sigma_1^2 \neq \sigma_2^2$$

$$Si \ \ F_{1-\frac{\alpha}{2},\, N_1-1, N_2-1} < \mathrm{F_o} < F_{\frac{\alpha}{2},\, N_1-1, N_2-1} \qquad se\ acepta\ H_o$$

para un contraste de dos colas. Debe tenerse en cuenta que la distribución F es positiva, no es simétrica y cumple que:

$$F_{1-\alpha,\, N_1-1, N_2-1} = \frac{1}{F_{\alpha,\, N_1-1, N_2-1}}$$



Problema 7: Una empresa de fertilizantes está estudiando dos proveedores de nitrato potásico. Se analizan 30 muestras de cada lote por emisión atómica. Se obtienen las siguientes purezas: 93.7 ± 0.7 % y 95.2 ± 0.4 %. Determina si los proveedores son similares.

Comprobaremos la homogeneidad de las varianzas:

$$\mathrm{F_o} = \frac{0.7^2}{0.4^2} = 3.06 > \mathrm{F_{0.025,29,29}} = 1.85 \qquad No\ son\ homogéneas$$

No podemos evaluar una varianza común ponderada. La varianza de la diferencia será:

$$s_d^2 = \frac{s_1^2}{N_1} + \frac{s_2^2}{N_2} = \frac{0.7^2}{30} + \frac{0.4^2}{30} = 0.0217 \quad s_d = 0.147$$

Como tenemos 30 medidas utilizaremos el estadístico de la normal que no necesita grados de libertad:

$$Z_o = \frac{95.2 - 93.7}{0.147} = 10.2 > z_{0.025} = 1.96 \qquad No\ son\ similares$$

El producto del segundo proveedor tiene mayor riqueza y será preferible

Problema 8: En un laboratorio se estudian el efecto del tiempo de tratamiento para la extracción de un analito en una muestra. Se investiga si el tiempo influye en los resultados y se obtuvieron los siguientes resultados en ppm:

30 minutos: 2.52, 2.47, 2.61, 2.56, 2.50

60 minutos: 2.57, 2.53, 2.51, 2.64

Se obtiene la media y la desviación estándar de cada tratamiento:

30 minutos $\bar{x}_{30} = 2.532$ $s_{30} = 0.0545$

60 minutos $\bar{x}_{60} = 2.563$ $s_{60} = 0.0574$

Comprobaremos la homogeneidad de las varianzas:

$$\mathrm{F_o} = \frac{0.0574^2}{0.0545^2} = 1.109 < \mathrm{F}_{0.025,3,4} = 9.98 \qquad Si\ son\ homogéneas$$

Evaluamos la varianza común ponderada:

$$s^2 = \frac{4 \times 0.0545^2 + 3 \times 0.0574^2}{5 + 4 - 2} = 0.00311 \quad s = 0.0558$$

Los grados de libertad son $v = 5 + 4 - 2 = 7$, y el estadístico será:

$$\mathrm{T_o} = \frac{2.563 - 2.532}{0.0558 \times \sqrt{\frac{1}{5} + \frac{1}{4}}} = 0.828 < t_{0.025,7} = 2.365 \qquad Son\ similares$$

La conclusión es que no mejora la extracción con más tiempo.

Presentación de resultados

Un resultado analítico no es útil si no se conoce su incertidumbre. El procedimiento ideal es indicar su intervalo de confianza. También se puede indicar la desviación estándar junto con el número de datos utilizado para calcularla. Una vez conocida la incertidumbre, es práctica común indicar el resultado con sus cifras significativas. Se aplican las siguientes reglas:

- Las cifras significativas son todas las que se conocen con certeza y el primer (o los dos primeros) dígitos inciertos. Los ceros a la izquierda no son significativos.
- es un uno) y es un indicativo de las cifras inciertas en el resultado lo que permite redondear su valor.
- El último cinco se redondea al número par más cercano.
- Es importante no efectuar ningún redondeo hasta la presentación final del resultado y utilizar al menos una cifra más de las significativas en los cálculos para evitar errores de redondeo. Los números redondeados no deben utilizarse en cálculos posteriores pues supondría cometer un error de redondeo que puede ser importante si la incertidumbre es alta.

Referencias

2002/657/EC. Decisión de la Comisión Europea de 12 de agosto de 2002 por la que se aplica la Directiva 96/23/CE del Consejo en cuanto al funcionamiento de los métodos analíticos y la interpretación de los resultados

ISO 5725. Accuracy (trueness and precision) of measurement methods and results.