

TEMA 3

INTRODUCCIÓN A LA QUIMIOMETRÍA

Cualquier medida llevada a cabo en el laboratorio analítico lleva asociado algún error. Los errores en los resultados de análisis bioanalíticos pueden tener efectos sociales y personales de importancia. El verdadero valor de algo es imposible medir, siendo la mejor opción la aplicación cuidadosa de una técnica establecida. Los científicos usan cálculos estadísticos para afinar sus juicios relativos a la calidad de las medidas experimentales. En este tema se tratan las aplicaciones más comunes de las pruebas estadísticas al tratamiento de datos analíticos.

1. La incertidumbre en las medidas científicas

- 1.1. Precisión y exactitud**
- 1.2. Tipos de error experimental. Propagación de errores**
- 1.3. Cifras significativas**

2. Estadística

- 2.1. Distribución de Gauss**
- 2.2. Intervalos de confianza**
- 2.3. Comparación de medias utilizando la t de Student**
- 2.4. Comparación de desviaciones estándar utilizando el test F**
- 2.5. Test Q de datos sospechosos**

1. LA INCERTIDUMBRE EN LAS MEDIDAS CIENTÍFICAS

En el trabajo científico es común trabajar tanto con números exactos como inexactos. Al escribir y manipular números tenemos que tener clara la diferencia entre los dos tipos de números. Como ejemplo de números exactos, tenemos los números enteros o fracciones (1, 2, 3...; $\frac{1}{2}$, $\frac{3}{4}$...), las constantes matemáticas (π , e ...) y las relaciones (1 kg=1000g; 4184 J=1 cal...). Los números inexactos son aquellos obtenidos a través de mediciones experimentales, siempre conllevan una incertidumbre asociada.

Como ya se indicó en el Tema 1, la definición de varias réplicas de la muestra resulta indispensable para estimar la fiabilidad de los resultados, ya que un único análisis no informa acerca de la variabilidad asociada. Ya que cada resultado tiene a ser distinto de los demás, de forma general se considera al valor central del conjunto como la mejor estimación. Supongamos que se determina la concentración de hierro en seis alícuotas de una misma muestra, cuyo valor verdadero de contenido de hierro es de 20,00 ppm. El valor central corresponde a 19,78 ppm y debe ser más fiable que cualquiera de los resultados individuales.

La medida de la tendencia central más utilizada es la **media**, también conocida como media aritmética, valor medio, promedio o como más adelante veremos media muestral. La media se obtiene al dividir la suma de los valores obtenidos en todas las medidas entre el número de medidas (N) del conjunto.

La **mediana** es el resultado que parece justo como valor central cuando los datos se ordenan en orden creciente o decreciente. Para un número de resultados impar, la mediana se obtiene de forma directa, si el nº de resultados es par, se calcula el valor medio del par central, que corresponderá a la mediana.

En casos ideales, la media y la mediana coinciden, pero suelen diferir cuando el número de medidas del conjunto es pequeño.

1.1. PRECISIÓN Y EXACTITUD

La **precisión** es una medida de la reproducibilidad de un resultado. Si se mide una cantidad varias veces exactamente de la misma manera y los valores obtenidos se aproximan mucho entre sí, se dice que la medida es precisa. Si los valores varían mucho entre sí, se dice que la medida no es precisa.

Son tres los términos de uso generalizado para describir la precisión de un conjunto de resultados: desviación estándar (expresada como s , e ó SD), varianza y coeficiente de variación (expresado como $\%s$, $\%e$, RSD ó DER):

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

$$\text{Varianza} = s^2$$

$$\%s = \frac{s}{x} 100$$

Las tres funciones anteriores informan de cuánto difiere de la media cada medida particular (x_i), lo que se conoce como desviación de la media (d_i):

$$d_i = |x_i - \bar{x}|$$

La relación entre d_i y los tres términos de precisión se detallará más adelante.

La **exactitud** describe la proximidad del valor medido respecto al valor “verdadero” o “aceptado”. Si se dispone de un estándar conocido, por ejemplo un material de referencia certificado, la medida será exacta si el valor obtenido es próximo al valor certificado.

Para comprender la diferencia entre precisión y exactitud veamos las siguientes situaciones:

- Una medida puede ser reproducible pero errónea. Por ejemplo, si se comete un error al preparar una disolución patrón de Fe, ésta no tendrá la concentración deseada. Al llevar a cabo la cuantificación de Fe en una muestra repetidas veces, los resultados pueden ser muy precisos pero inexactos, porque la concentración real de la disolución patrón no es la que deseábamos preparar. En definitiva: buena precisión, mala exactitud.
- Pero también puede ocurrir que las medidas sean poco reproducibles, pero en torno al valor correcto, porque la disolución patrón fuese preparada sin errores pero el método analítico empleado no sea muy reproducible. En definitiva: mala precisión, buena exactitud.
- Situación ideal: procedimientos exactos y precisos.

La exactitud es con frecuencia más difícil de determinar que la precisión, pues para la precisión basta con analizar varias réplicas de la muestra. Pero para determinar la exactitud se requiere el conocimiento del valor verdadero.

Para obtener el valor verdadero de un parámetro, éste habrá tenido que ser medido experimentalmente y, como ya sabemos, toda medida experimental lleva asociada un error. Podríamos definir el valor verdadero como el obtenido por una persona experimentada empleando un procedimiento bien establecido o, mejor aún sería preferible que ese valor hubiese sido obtenido a través de diferentes procedimientos analíticos y en distintos laboratorios. En cualquier caso, el error

asociado podría minimizarse pero nunca anularse, por ello parece más apropiado hablar de valor aceptado más que verdadero.

La exactitud se expresa en términos del error absoluto o relativo.

Error absoluto:

$$\text{Error absoluto (E)} = \text{Valor verdadero} - \text{Valor obtenido}$$

Considerando que el valor verdadero de un parámetro es igual a 0,2 ppm. El error absoluto al obtener 19,8 ppm sería $-0,2$ ppm. El error absoluto del resultado 20,1 ppm es $+0,1$ ppm. Obsérvese que se mantiene el signo al expresar el error, pues nos informa si se produce por exceso o defecto.

Error relativo: Frecuentemente este parámetro es más útil que el error absoluto. El error relativo para el caso concreto citado corresponde a -1% .

1.2. TIPOS DE ERROR EXPERIMENTAL

Cualquier medida lleva asociada una incertidumbre, que se llama error experimental. Los resultados pueden expresarse con un mayor o menor grado de confianza, pero nunca con total certeza. Los análisis químicos se ven afectados al menos por dos tipos de errores: SISTEMÁTICOS y ALEATORIOS. Existe otro tipo de error denominado BRUTO.

ERROR SISTEMÁTICO: hace que la media de un conjunto de datos difiera del valor aceptado. El error sistemático, también llamado error determinado, se origina principalmente por un fallo del diseño del experimento o por un fallo del equipo. Si se repite el experimento en idénticas condiciones, vuelve a producirse este error; provocando que todos los resultados se hallen por encima o por debajo del valor aceptado. Este tipo de error es difícil de descubrir aunque no imposible.

Supongamos que para la adición de reactivo valorante en una volumetría directa utilizamos una bureta de 10 mL, que no hemos calibrado previamente. El volumen liberado hasta alcanzar el punto final de la valoración es de $8,62 \pm 0,02$ mL. Si supuestamente se está obteniendo una concentración de analito por exceso, es probable que el volumen real liberado de valorante sea superior a los valores de lectura de bureta. La calibración del material volumétrico sería una buena forma de corregir este error sistemático.

Existen diferentes vías para detectar un error sistemático, algunas de ellas son:

- a- Analizar muestras de composición conocida, tales como materiales de referencia certificados. El método ensayado debe reproducir el resultado certificado.
- b- Analizar muestras blanco (que no contengan el analito). Si se observa un resultado distinto de cero, el método acarrea un error por exceso.
- c- Usar métodos analíticos diferentes para llevar a cabo el análisis. Si los resultados no concuerdan, hay un error en uno o más de los métodos.
- d- Comparación entre varios laboratorios. Designamos distintas personas para analizar la misma muestra mediante el mismo método o distintos métodos analíticos.

El ERROR ALEATORIO, también llamado error indeterminado, se origina por efecto de variables incontroladas. Tiene igual probabilidad de ser positivo que negativo. Siempre está presente y no puede ser corregido.

Como ejemplo de error aleatorio podemos citar el producido al tomar la lectura de volumen en una bureta. Si una determinada lectura es llevada a cabo por personas diferentes, con toda probabilidad cada persona daría un valor distinto, pues la interpolación entre rayas es algo subjetiva. Incluso una misma persona al leer la misma magnitud varias veces puede dar lecturas distintas. Otra causa de error aleatorio es el ruido eléctrico de un instrumento, que se presenta como fluctuaciones tanto positivas como negativas con similar frecuencia y no pueden ser eliminadas completamente. Aunque estos errores no pueden eliminarse, sí se pueden minimizar mejorando el trabajo experimental.

Un tercer tipo de error es el conocido como ERROR BRUTO. Generalmente ocurre de forma ocasional y suele ser grande. Puede ser común que la persona cometa este tipo de errores, por ejemplo perder parte del precipitado antes de su pesada en un análisis gravimétrico. Los errores brutos dan lugar a valores atípicos, son valores que difieren mucho de los demás en un conjunto de datos de medidas replicadas. La localización de un valor atípico puede llevarse a cabo a través de pruebas estadísticas.

PROPAGACIÓN DE ERRORES

En la mayoría de los trabajos experimentales es necesario hacer operaciones aritméticas con varios números, teniendo cada uno de ellos un error aleatorio. La incertidumbre más probable del resultado no es simplemente la suma de los errores individuales, pues es probable que algunos de ellos sean positivos y otros negativos. Por esta razón, es probable que algunos de los errores se anulen entre sí.

En las operaciones de suma y resta para calcular la incertidumbre total se utiliza la incertidumbre o error absoluto de los términos individuales de acuerdo con la ecuación siguiente:

$$s_4 = \sqrt{s_1^2 + s_2^2 + s_3^2}$$

donde s_1 , s_2 y s_3 son las desviaciones estándar de los términos de la operación (en este caso concreto tres términos) y s_4 es la desviación estándar del resultado de la operación.

En las operaciones de multiplicación y división, lo primero que hay que hacer es expresar todas las incertidumbres relativas en porcentaje. Seguidamente el error del producto o cociente se calcula según la ecuación siguiente:

$$\%s_4 = \sqrt{(\%s_1)^2 + (\%s_2)^2 + (\%s_3)^2}$$

seguidamente a partir de $\%s_4$ se puede obtener s_4 .

1.3. CIFRAS SIGNIFICATIVAS

El número de cifras significativas es el mínimo número de dígitos necesarios para escribir un valor dado en notación científica sin perder exactitud. Otra forma de definir lo que son las cifras significativas de una determinada medida es el número de dígitos conocidos con certeza más el primero incierto. Por ejemplo: al tomar la lectura de volumen en una bureta de 50 mL donde el nivel de líquido se halla entre 9,6 y 9,7 mL, si no coincide ni con 9,60 ni con 9,70 de forma exacta, es necesario estimar el valor de la segunda cifra decimal. Suponiendo que se estima 0,008, el volumen medido sería de 9,68 mL, que tiene 3 cifras significativas.

El cero puede ser significativo o no dependiendo de su ubicación en el número:

- Un cero rodeado por otros dígitos siempre es significativo: 30,22 mL
- Los ceros que sitúan sólo la coma decimal no son significativos: 0,03022 L, sigue teniendo 4 cifras significativas.
- Los ceros al final del número pueden ser significativos o no: Si se expresa el volumen de un vaso como 2,0 L, la presencia del cero indica que el volumen se conoce hasta unas décimas de litro, por lo que ese 0 es significativo. Si este mismo volumen lo expresamos como 2000 mL, si no se conoce el volumen en centésimas de litro, seguimos teniendo dos cifras significativas. Otro ejemplo: El número de cifras significativas en 92500 es ambiguo, para detallarlo tendría que poner:
 - $9,25 \times 10^4$ tres cifras significativas
 - $9,250 \times 10^4$ cuatro cifras significativas
 - $9,2500 \times 10^4$ cinco cifras significativas

Habría que indicar sólo uno de estos tres números.

Se plantea ahora la cuestión de cuántos dígitos deben mantenerse en un resultado tras hacer operaciones aritméticas:

En la suma y resta si los términos tienen el mismo número de dígitos, el resultado tendrá el mismo número de decimales que los términos individuales de la operación. Si los términos que se suman o restan no tiene el mismo número de cifras significativas, el resultado debe expresarse con el mismo número de cifras decimales que la magnitud con menos cifras decimales. Ahora bien, debe operarse con todas las cifras decimales y redondear al final. Cuando se opera con números expresados en notación científica, todos los términos deben de tener el mismo exponente.

En la multiplicación y división, el resultado debe expresarse con el mismo número de cifras significativas del factor con menos cifras significativas. Las potencias de diez no influyen en el número de cifras que se pueden mantener en el resultado final.

2. ESTADÍSTICA

La estadística nos proporciona herramientas para aceptar conclusiones que tienen alta probabilidad de ser correctas y rechazar las conclusiones cuya probabilidad de ser incorrectas es alta.

2.1. DISTRIBUCIÓN DE GAUSS

Cuando un experimento se repite un número elevado de veces y los errores son solamente aleatorios, los resultados tienden a agruparse simétricamente en torno al valor medio, asemejándose la representación la agrupación de los resultados a una curva ideal llamada Distribución de Gauss. Aunque en realidad en un laboratorio no llevamos a cabo un análisis un número infinito de veces, ni siquiera medimos una misma muestra 100 ó 200 veces, lo más habitual es repetirlo de 2 a 5 veces, también podemos estimar los parámetros estadísticos que caracterizan a una serie de un número grande de medidas.

Decimos que la variación de los datos experimentales está distribuida normalmente cuando al repetir medidas aparece una distribución en forma de campana. Existe la misma probabilidad de que una medida sea mayor o menor que la media. Además, la probabilidad de observar un valor disminuye a medida que aumenta la distancia a la media.

La media constituye el centro de la distribución y la desviación estándar (s) mide el ancho de la distribución. Por tanto, puede definirse la desviación estándar como una medida del grado de proximidad de los datos en torno al valor de la media. Cuanto menor es s , más estrechamente se agrupan los datos alrededor de la media y decimos que la precisión es alta.

Para una serie infinita de datos, la media se designa con la letra minúscula griega “ μ ” y la desviación estándar con la letra minúscula griega “ σ ”. Hay que tener claro que nunca medimos μ y σ , pero los valores de la media y la desviación estándar se acercan a μ y σ , respectivamente, a medida que aumenta el número de medidas.

Cuanto mayor es la desviación estándar, más ancha es la curva de Gauss. En toda curva de Gauss, el área comprendida en el intervalo, desde el resultado de restar s al valor medio y el resultado de sumar s al valor medio, supone el 68,3% del área total. Es decir, es de esperar que más de 2/3 de las medidas no disten de la media en más de una desviación estándar. Además el 95,5% del área está entre el resultado de restar al valor medio el doble de la desviación estándar y el resultado de sumar al valor medio el doble de la desviación estándar. El 99,7% del área se encuentra entre los extremos correspondientes a restar al valor medio $3s$ y sumar al valor medio $3s$. Supongamos que se emplean dos métodos analíticos diferentes, A y B para llevar a cabo la determinación de hierro en una muestra de sangre, y que el método A tiene una RSD de 0,4% y B de 1,1%. Se puede esperar que aproximadamente 2/3 de las medidas obtenidas mediante el método A estén dentro del 0,4% de la media, mientras que para el método B, 2/3 de las medidas estarán dentro del 1,1% de la media.

La ecuación de la curva de Gauss es la siguiente:

$$y = \frac{1}{\sigma \sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

Donde e es la base de los logaritmos naturales y μ y σ la media verdadera y su desviación estándar asociada. Para una serie finita de datos, μ se aproxima al valor medio y σ a la desviación estándar. Resulta muy útil expresar las desviaciones respecto de la media como múltiplos de la desviación estándar. Esto se lleva a cabo transformando el valor de x en otro valor numérico llamado z a través de la siguiente expresión:

$$z = \frac{x - \mu}{\sigma} \approx \frac{x - \bar{x}}{s}$$

De este modo, la probabilidad de medir z en un cierto intervalo es igual al área de ese intervalo. Dado que la suma de las probabilidades de todas las medidas es la unidad, el área debajo de la curva desde $-\infty$ a $+\infty$ debe ser la unidad.

Los valores de las áreas debajo de cada porción de la curva de Gauss se encuentran tabulados.

2.2. INTERVALOS DE CONFIANZA

La t de Student se usa muy frecuentemente para expresar intervalos de confianza y para comparar resultados de diferentes experimentos. Esta herramienta se podría utilizar para calcular la probabilidad de que el recuento de glóbulos rojos en un paciente se encuentre dentro del intervalo “normal”.

¿Cómo se calculan intervalos de confianza?

Si se dispone de un número limitado de medidas, que es la situación habitual en análisis químico, no podemos hallar la verdadera media de la población (μ) ni la verdadera desviación estándar (σ). Lo que podemos determinar es la media muestral y la desviación estándar muestral (s). El intervalo de confianza es una expresión que nos informa de que la verdadera media μ , está probablemente a una cierta distancia de la media medida. El intervalo de confianza de μ viene dado por la siguiente ecuación

$$\mu = \bar{x} \pm \frac{t s}{\sqrt{n}}$$

donde s es la desviación estándar muestral (la obtenida experimentalmente), n el número de observaciones y t es la t de Student tabulada para $(n-1)$ grados de libertad. A partir de la ecuación del intervalo de confianza se observa claramente que se puede reducir la incertidumbre aumentando el número de análisis (n).

Se presenta un ejercicio para clarificar el cálculo de intervalos de confianza: La determinación del contenido de proteínas en una muestra de sangre aporta los siguientes resultados replicados: 120,6; 118,9; 131,0; 124,7 y 126,9 mg/dL. Al calcular el intervalo de confianza al 50% de nivel de confianza del contenido de proteínas se obtienen como extremos de dicho intervalo: 12,4 y 12,6 mg/dL. ¿Qué significa el resultado obtenido? Que existe un 50% de probabilidad de que el valor de la verdadera media (μ) sea mayor de 12,4 y menor de 12,6.

Si se incrementa el porcentaje del nivel de confianza, ¿cómo afectará a la amplitud del nuevo intervalo de confianza? Si se incrementa el nivel de confianza la amplitud del intervalo también lo hará. Lógicamente si queremos aumentar la probabilidad de

obtener la verdadera media, al ampliar el intervalo aumentará la probabilidad de contener dicho valor. Esto es obvio dado que el valor de la t de Student aumenta para un mismo número de grados de libertad con el nivel de confianza, y se encuentra en el numerador de la expresión del cálculo del intervalo de confianza.

El número de medidas también influye en la amplitud del intervalo de confianza. Cuantas más veces se mide una cantidad, más confianza se tiene que el valor de la media de las medidas hechas esté próximo a la verdadera media de la población n . De hecho, la incertidumbre disminuye en proporción a $1/(n)^{1/2}$, siendo n el número de medidas.

Si mantenemos constante el número de medidas y el nivel de confianza, se cumplirá que a mayor desviación estándar mayor será la amplitud del intervalo de confianza obtenido.

2.3. COMPARACIÓN DE MEDIAS UTILIZANDO LA t DE STUDENT

El test t se utiliza para comparar dos conjuntos de medidas y poder afirmar si son o no diferentes. A nivel estadístico se trata de comprobar la “hipótesis nula”, o sea se parte de afirmar que los valores medios de dos series de medidas no son diferentes. Idénticos no pueden ser, puesto que los errores aleatorios son inevitables. De manera que la Estadística predice una probabilidad de que la diferencia entre las dos medidas pueda deberse a dichos errores aleatorios. Si hay menos de un 5% de probabilidad de que la diferencia se deba a errores aleatorios, se suele rechazar la hipótesis nula. Si existe más de un 95% de probabilidad de que la diferencia se deba solo a errores aleatorios, se acepta la hipótesis nula, concluyendo que los dos grupos de medidas no son significativamente diferentes, en el caso de que los cálculos se desarrollen al 95% de nivel de confianza.

En la comparación de dos conjuntos pueden darse tres casos, que se tratan de forma algo diferente:

CASO 1: Se mide una cantidad varias veces obteniéndose un valor medio y una desviación estándar. Queremos comparar el resultado con un resultado conocido y aceptado. La media no coincide exactamente con el resultado aceptado. Será necesario responder a la pregunta: ¿Coincide el resultado medido con el resultado conocido dentro del error experimental?

CASO 2: Se mide una cantidad varias veces usando dos métodos diferentes, obteniéndose dos resultados distintos, cada uno con su valor medio y su

desviación estándar. Será necesario responder a la pregunta: ¿Concuerdan entre sí los dos resultados dentro del error experimental?

CASO 3: Se mide una vez la muestra 1 mediante el método A y otra vez con el método B, y no dan el mismo resultado. Lo mismo se hace con la muestra 2, y de nuevo los resultados no son iguales. Se repite el proceso con n muestras diferentes. La pregunta es: ¿Concuerdan los dos métodos dentro del error experimental o son sistemáticamente diferentes?

Para la resolución de los tres casos puede aplicarse un patrón común que consiste en las siguientes etapas:

1. Selección del caso, de acuerdo con los datos proporcionados.
2. Obtención del valor de t calculado aplicando las fórmulas matemáticas correspondientes al caso seleccionado previamente.
3. Obtención del valor de t tabulado según cada caso particular.
4. Comparación de los valores de t calculada y tabulada.

CASO 1: COMPARACIÓN DE UN RESULTADO MEDIDO CON UN VALOR CONOCIDO.

Supongamos el caso concreto en el que se compró una muestra de orina sintética que correspondía a un material estándar de referencia certificado por el NIST (National Institute Standard Technology) con un contenido en selenio de $(30,1 \pm 0,9)$ ng/mL. Se pretende validar un nuevo método analítico para la determinación de selenio en muestras de orina, para ello se analiza dicho material de referencia mediante el método a validar. Siendo los valores obtenidos para 4 alícuotas del material de referencia 27,6; 29,3; 28,1 y 28,9 ng/mL. Dado que el valor medio de estas cuatro medidas (28,5 ng/mL) no coincide exactamente con el valor certificado (valor verdadero o aceptado, en este caso 30,1 ng/mL), es necesario aplicar las fórmulas correspondientes a este caso para calcular el valor de t. Ya que el valor de t_{tabulada} (3,182) es menor que el de t calculada, podemos afirmar que para este caso concreto existen diferencias significativas al nivel de confianza del 95% entre los dos resultados (el obtenido y el certificado), o lo que es lo mismo la probabilidad de que dichos valores sean iguales es inferior al 5%.

Los test estadísticos no eximen de tener que tomar personalmente la última decisión de aceptar o rechazar una conclusión. Sólo son una guía en términos de probabilidad. Por ejemplo, según los resultados obtenidos en el problema informamos de que el valor de los resultados del análisis de selenio es diferente del valor conocido.

Sin embargo, con sólo 4 medidas, sería razonable repetir el análisis varias veces más para tener más seguridad en dicha afirmación.

CASO 2: COMPARACIÓN DE MEDIDAS REPLICADAS

Se propone ahora el caso en el que se mide el contenido de selenio en una única muestra de orina, mediante dos métodos analíticos diferentes, por ejemplo espectrometría de absorción atómica con atomización electrotérmica (ETAAS) y espectrometría de absorción atómica con generación de hidruros (HG-AAS). Supongamos que se miden cuatro alícuotas de la muestra con cada técnica. Una vez calculados el valor medio y la desviación estándar para cada grupo de medidas, se plantea la pregunta ¿Concuerdan entre sí los dos resultados obtenidos dentro del error experimental? Para responder a esta pregunta se calcula $s_{combinada}$ mediante la ecuación correspondiente y seguidamente el valor de $t_{calculada}$, que se comparará con el valor tabulado de t para $(n_1 + n_2 - 2)$ grados de libertad. En el caso propuesto el número de medidas llevadas a cabo con cada método es el mismo, $n_1 = n_2$, pero podrían ser diferentes. Si por ejemplo se obtiene que $t_{calculada} > t_{tabulada}$, afirmaremos que existen diferencias significativas al nivel de confianza del 95% entre los resultados obtenidos mediante los dos métodos.

CASO 3: COMPARACIÓN DE PARES DE MEDIDAS

En este caso se trata de dos métodos analíticos diferentes con los que se hace una única medida usando muestras diferentes. No se duplica ninguna medida. ¿Son sistemáticamente diferentes ambos métodos?. Aplicamos el test de las diferencias individuales entre los resultados de cada muestra, calculando el valor de t y comparándolo con el tabulado. Si $t_{calculada} > t_{tabulada}$ (al 95% de nivel de confianza) afirmamos que entre las dos técnicas existen diferencias significativas al nivel de confianza citado.

2.4. COMPARACIÓN DE DESVIACIONES ESTÁNDAR CON EL TEST F

El test F nos informa si dos desviaciones estándar son significativamente diferentes entre sí. F es el cociente de los cuadrados de las desviaciones estándar. Se coloca siempre en el numerador la desviación estándar mayor, de modo que se cumpla que el valor calculado para F sea mayor o igual a 1. El valor de F se halla tabulado para un nivel de confianza del 95%. Si $F_{calculada} > F_{tabulada}$, se concluye que los dos métodos comparados sí presentan desviaciones estándar significativamente diferentes al nivel de confianza del 95%.

2.5. TEST Q DE DATOS SOSPECHOSOS

Cuando un dato no es coherente con los restantes, es decir, es muy alto o muy bajo respecto del resto de datos, se puede usar el test Q como ayuda para decidir si se mantiene o se desecha dicho dato sospechoso. Consideremos los 5 resultados siguientes: 130,1; 130,7; 128,8; 137,8 y 131,4. ¿Se rechaza o se mantiene el valor 137,8?. Para aplicar el test Q, se ordenan los datos en orden creciente y se calcula Q, según la ecuación:

$$Q_{\text{calculada}} = \text{divergencia/recorrido}$$

donde el recorrido es la dispersión máxima entre los datos y la divergencia es la diferencia entre el valor sospechoso y el valor más próximo.

Los valores de Q también se hallan tabulados para el 90% de nivel de confianza, y se selecciona aquel valor que corresponda al número de medidas llevadas a cabo, no al número de grados de libertad. Si $Q_{\text{calculada}} > Q_{\text{tabulada}}$, se descarta el punto sospechoso. En el caso propuesto el dato sospechoso no debe descartarse.

En realidad la decisión final depende de uno mismo, pues hay quien afirma que no se debe descartar nunca un dato a no ser que se sepa que existe un error de procedimiento que condujo a esa medida particular; hay quien repite la medida sospechosa varias veces hasta asegurarse si está o no realmente fuera de los esperable.