



## QUÍMICA ANALÍTICA - CAPÍTULO 1 TRATAMIENTO DE DATOS ANALÍTICOS

### 1. Introducción

Toda medida física está sujeta a un grado de incertidumbre que, en el mejor de los casos, puede ser reducido a un valor aceptable, pero nunca podrá eliminarse totalmente. Determinar la magnitud de esta incertidumbre puede ser difícil y requiere de un esfuerzo adicional al de la medición y de criterio por parte del observador. Sin embargo, cualquier determinación cuantitativa en la que se exprese un resultado numérico sin el conocimiento de la incertidumbre asociada con él es un dato inútil. Por otra parte, un resultado de menor exactitud podrá ser de utilidad si se conoce el límite de error que lo afecta. Es decir, *no existen resultados cuantitativos válidos si no van acompañados de una estimación de los errores inherentes a ellos*. Veamos el siguiente ejemplo sencillo. Supongamos que un químico sintetiza un compuesto creyendo que es completamente nuevo, estudia una propiedad física arbitraria y el compuesto arroja un valor de 104 (en unidades arbitrarias). La literatura, sin embargo, indica que ningún compuesto descubierto hasta el momento ha arrojado un valor de más de 100 cuando se estudia en las mismas condiciones experimentales. La pregunta que surge naturalmente es si el químico descubrió un compuesto nuevo. Por supuesto que la respuesta a esta pregunta reside en el *grado de confianza* que podemos asignarle al valor experimental de 104, es decir, qué errores están asociados con el dato. Si el análisis posterior indica que el resultado es correcto dentro de 2 unidades (arbitrarias), es decir que el valor verdadero se encuentra dentro de  $104 \pm 2$ , entonces es posible que se haya caracterizado un nuevo material. Pero si las estimaciones de error posteriores muestran que el error puede alcanzar a 10 unidades ( $104 \pm 10$ ), entonces es muy probable que el verdadero valor sea en realidad menor que 100, en cuyo caso, no se sintetizó un nuevo compuesto. De este ejemplo sencillo se deduce que *es esencial conocer los errores experimentales para interpretar adecuadamente el resultado obtenido*.

### 2. Definiciones de parámetros estadísticos a usar

Las determinaciones de la Química Analítica se basan mediciones de volúmenes, masas, magnitudes eléctricas, etc. con el objetivo de encontrar el "verdadero valor,  $\mu_0$ " de contenido de un analito en una muestra. Como no existen formas de alcanzar este valor verdadero, en la práctica el químico realiza un número  $n$  de análisis independientes sobre porciones individuales de muestra, elige el mejor valor para la variable y estima la incertidumbre asociada al mismo. El mejor valor de la serie se expresa a partir de la media:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad [1]$$

que es un parámetro de posición de la distribución de datos. Por otra parte, las *variaciones* entre los resultados individuales proporcionarán una estimación del "margen de confianza" que se puede asignar a éste resultado. Para esto se calcula la varianza de la serie de datos respecto de la media:

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

o bien, la desviación estándar de la serie de  $n$  experimentos individuales,

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \quad [3]$$

Un parámetro que se emplea con fines comparativos de la dispersión en los datos es el cociente entre la desviación estándar y la media de los datos, esto es, la desviación estándar relativa (RSD), denominada también "coeficiente de variación porcentual",

$$CV \quad RSD\% = 100 \times \frac{s}{\bar{x}} \quad [4]$$

Para un conjunto de medias obtenidas sobre la misma muestra, la desviación estándar de la media se calcula como:

$$s_{\bar{x}} = \frac{s}{\sqrt{n}} \quad [5]$$

### 3. Precisión y Exactitud

El término *precisión* describe la repetibilidad de resultados obtenidos en forma idéntica. Puede definirse como la concordancia entre los valores numéricos de dos o más mediciones que se obtuvieron en forma idéntica. La varianza, desviación estándar y desviación estándar relativa son estimaciones simples de la precisión de un análisis.

El término *exactitud* indica la proximidad de una medida a su valor aceptado como "verdadero",  $\mu_0$ . Vale aclarar la distinción fundamental entre precisión y exactitud: exactitud supone una comparación con un valor verdadero o aceptado como tal, mientras que la precisión indica la concordancia entre las medidas realizadas de la misma forma.

La exactitud se expresa en términos de error. El error absoluto,  $\epsilon$ , de una medida es la diferencia entre el valor informado,  $\bar{x}$ , y el valor verdadero  $\mu_0$ :

$$\epsilon = \bar{x} - \mu_0 \quad [6]$$

En el caso del error, el signo asociado con el error es tan importante como el valor mismo, porque el analista deberá saber si el efecto del error ha provocado un aumento o una *disminución en el resultado informado*.

Finalmente el error relativo y error relativo porcentual expresan una medida de exactitud más útil que el error absoluto. Se definen como:

$$\epsilon_r = \frac{(\bar{x} - \mu_0)}{\mu_0} = \frac{\epsilon}{\mu_0} \quad [7a] \quad \text{y} \quad \epsilon_r\% = 100 \times \frac{(\bar{x} - \mu_0)}{\mu_0} \quad [7b]$$

Supongamos el siguiente ejemplo. Un estudiante realiza una titulación de 10,00 mL de NaOH *exactamente* 0,1 M. (un milimol de base) con HCl de igual molaridad (asumimos que las concentración de ácido y la concentración y el volumen de NaOH son datos que se conocen "sin error"). Los resultados obtenidos sobre cinco porciones de muestra son los indicados en la segunda columna de la tabla:

Muestra	Volumen HCl (mL)	Desvío de la media (mL), $ x_i - \bar{x} $	$(x_i - \bar{x})^2$
$x_1$	10,08	0,02	0,0004
$x_2$	10,11	0,01	0,0001
$x_3$	10,09	0,01	0,0001
$x_4$	10,10	0	0
$x_5$	10,12	0,02	0,0004
	$\bar{x} = 10,10$	$\bar{d} = 0,012$	$\Sigma = 0,001$

El promedio de las lecturas de volumen es de 10,10 mL, el desvío medio de 0,012 mL, la desviación estándar es  $s = \sqrt{\frac{0,001}{5-1}} = 0,016$  y la desviación estándar relativa es  $RSD\% = (0,016/10,10) \times 100 = 0,16\%$ .

En este caso en particular, la desviación estándar relativa es pequeña (0,16%), lo que indica que el resultado de volumen obtenido (10,10 mL) es preciso.

El cálculo del error cometido en la determinación supone el conocimiento del verdadero valor de volumen de HCl. En este ejemplo hipotético de titulación asumimos que el Erlenmeyer contiene 1 milimol de base (dato que se conoce *sin error*), por lo que sería de esperar que el volumen de HCl a gastar fuese de exactamente 10,00 mL. En consecuencia, el error absoluto y relativo cometido es:

$$\epsilon = 10,10 - 10,00 = 0,10 \text{ mL} \quad \text{y} \quad \epsilon_r\% = 100 \times (0,10 / 10,00) = 1\%$$

Exactitudes aceptables en un análisis volumétrico usando indicadores son del orden de 0,1 – 0,2%. El error relativo obtenido por el estudiante indica que el resultado no es exacto. A partir de este resultado se deberá buscar y corregir la fuente de error que determina que los resultados obtenidos sean más altos que los valores esperados.

Obviamente, este ejemplo es completamente "hipotético". En un análisis cuantitativo (en este caso en una titulación), el analista desconoce de antemano el resultado de la medición. Efectuando el análisis sobre  $n$  porciones de muestra en forma idéntica será relativamente sencillo determinar la

precisión del resultado. No puede decirse lo mismo para la exactitud ya que para ello se requiere conocer el valor verdadero, es decir, la misma información que se desea obtener.

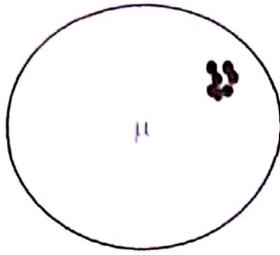
#### 4. Tipos de errores

Existe una distinción entre tres tipos de errores: groseros, aleatorios y sistemáticos.

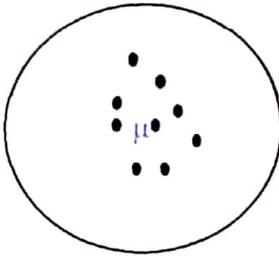
- a) **Errores groseros.** Los errores groseros se describen fácilmente: son errores tan graves que no queda otra alternativa que abandonar el experimento y empezar de nuevo. Como ejemplo puede citarse la avería de un instrumento, el derramamiento accidental de muestra, o descubrir que un reactivo que se suponía puro, en realidad estaba contaminado.
- b) **Errores sistemáticos** (o determinados) son aquellos a los que pueden atribuírsele una causa definida, los cuales en principio pueden ser calculados y tenidos en cuenta. Tales errores son caracterizados por ser unidireccionales. La magnitud de tales errores puede ser: a) constante, b) proporcional al tamaño de muestra o, c) variar con el tamaño de muestra en una forma compleja.
  - Errores sistemáticos constantes. Un ejemplo de error sistemático constante y de signo negativo es el causado por pérdida por solubilidad del precipitado en una determinación gravimétrica. Asumiendo que se pierden por solubilidad durante el lavado del precipitado una masa de precipitado de 0,5 mg, el error absoluto será constante (de -0,5 mg) y si la masa de sólido a pesar es de 500 mg, el error relativo será  $-0,5/500 = -0,1\%$ . Si se pesan, en cambio, 100 mg de precipitado el error será de -0,5%. Otro ejemplo de errores sistemáticos constantes son los que afectan a una titulación empleando un indicador, en este caso se requiere una cantidad mínima de reactivo titulante para dar lugar a un cambio de color. Este volumen, generalmente pequeño, es el mismo cualquiera sea el volumen total de reactivo titulante. Nuevamente, el error relativo aumentará a medida que el volumen de titulante sea menor. Es evidente que una forma de minimizar la influencia de este tipo de errores consiste en usar una muestra tan grande como lo permita el método.
  - Errores proporcionales al tamaño de muestra. Son errores sistemáticos cuyo valor absoluto se incrementa al aumentar la muestra. Un ejemplo es la presencia, en la muestra, de un contaminante que interfiere en el análisis si no se elimina previamente. Como ejemplo supongamos que se quiere determinar sulfato por gravimetría como sulfato de bario y la muestra contiene además anión fluoruro. Si no se elimina antes de la precipitación, el fluoruro también precipitará con cloruro de bario, dando una masa mayor que la real. Si se incrementa el tamaño de la muestra, el error absoluto aumentaría en forma proporcional (y el error relativo será constante).
  - Errores que varían en forma más compleja. Un ejemplo es el error cometido en la pesada de una sustancia higroscópica. Este error será siempre de signo positivo, aumentará con la superficie de muestra expuesta y dependerá finalmente del tiempo requerido para efectuar la pesada, de la humedad y de la temperatura ambiente.
- c) **Errores aleatorios** (o indeterminados) son el resultado de limitaciones naturales para realizar mediciones físicas. Siempre que se efectúa la lectura de una medición en un instrumento, los resultados obtenidos (en igualdad de condiciones) no son constantes y fluctúan alrededor de un valor medio. El origen de tales fluctuaciones no pueden identificarse porque son la resultante de innumerables incertidumbres individuales pequeñas e indetectables. Como indica su nombre, el signo de este error podrá ser a veces positivo, y a veces negativo. Existe siempre, no puede ser corregido y es el limitante de las determinaciones experimentales.

Para ilustrar el concepto de errores sistemáticos y aleatorios retomemos el ejemplo de la titulación de 10,00 mL de NaOH 0,1 M con HCl de igual molaridad. Además del estudiante anterior (estudiante A), la misma titulación (con los mismos reactivos, y empleando la misma bureta y pipeta aforada) la efectuaron otros 3 estudiantes. Los resultados fueron los siguientes:

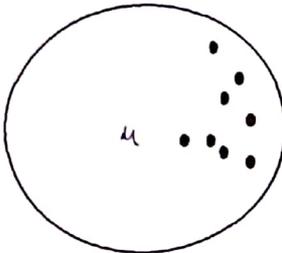
Resultados (mL)			
Estudiante A	Estudiante B	Estudiante C	Estudiante D
10,08	9,88	10,19	10,04
10,11	10,14	9,79	9,98
10,09	10,02	9,69	10,02
10,10	9,80	10,05	9,97
10,12	10,21	9,78	10,04
$\bar{x} = 10,10$	$\bar{x} = 10,01$	$\bar{x} = 9,90$	$\bar{x} = 10,01$
$s = 0,02$	$s = 0,17$	$s = 0,21$	$s = 0,03$
$\checkmark$ RSD% = 0,2%	RSD% = 1,7%	RSD% = 2,1%	RSD% = 0,33%
$\epsilon_r\% = +1\%$	$\epsilon_r\% = +0,1\%$	$\epsilon_r\% = -1\%$	$\epsilon_r\% = +0,1\%$
Preciso pero inexacto	Exacto e impreciso	Inexacto e impreciso	Exacto y preciso



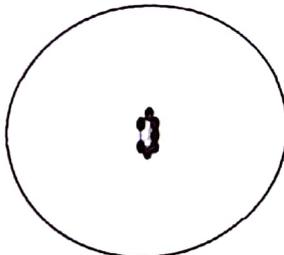
Los resultados obtenidos por el estudiante A son todos muy próximos entre sí, es decir son repetitivos (RSD= 0,2 %). Sin embargo, como ya se discutió, son *todos demasiado altos*. En este experimento "conocemos" que la respuesta correcta debe ser 10,00 mL. Resulta evidente que en el experimento desarrollado por el estudiante A han surgido dos tipos de errores: aleatorios, los cuales provocan que los datos individuales fluctúen alrededor del valor medio (10,10 mL) y sistemáticos, los cuales provocan que todos los resultados sean erróneos en el mismo sentido (todos son demasiado altos). Los errores de tipo aleatorio son los que afectan la precisión o repetibilidad o reproducibilidad de un experimento. Los errores sistemáticos en cambio inciden en la exactitud del resultado obtenido. La titulación realizada por el estudiante A deberá calificarse de precisa e inexacta.



El estudiante B ha obtenido resultados diferentes. El promedio de los cinco volúmenes (10,01 mL) es muy próximo al valor verdadero, por lo que puede calificarse como datos exactos (sin errores sistemáticos importantes). Pero la dispersión de resultados es grande, lo que indica una precisión insatisfactoria y la presencia de errores aleatorios sustanciales. La comparación de estos resultados con los obtenidos por el estudiante A demuestra claramente que los errores aleatorios y sistemáticos son independientes uno del otro.



El trabajo del estudiante C no es ni preciso (intervalo de datos de 9,69 a 10,19 mL, desvío estándar relativo: 2,1%) ni exacto ( $\bar{x}$  = 9,90 mL).



El estudiante D ha logrado a la vez resultados precisos (intervalo de 9,97 a 10,04 mL, RSD%=0,28%) y exactos ( $\bar{x}$  = 10,01 mL).

En síntesis, la precisión de un análisis se puede calcular a partir de repeticiones del mismo. La exactitud demanda el conocimiento del valor verdadero,  $\mu_0$ , que nunca se conoce. Existen algunas alternativas para estimar la exactitud de un método: una es el análisis de muestras de materiales certificados (muestras donde la concentración de analito se conoce con alto grado de certeza), o bien realizando el análisis por un método completamente distinto y cuya precisión y exactitud sean conocidas.

### 5. Curva normal de error. Intervalo de confianza de la media

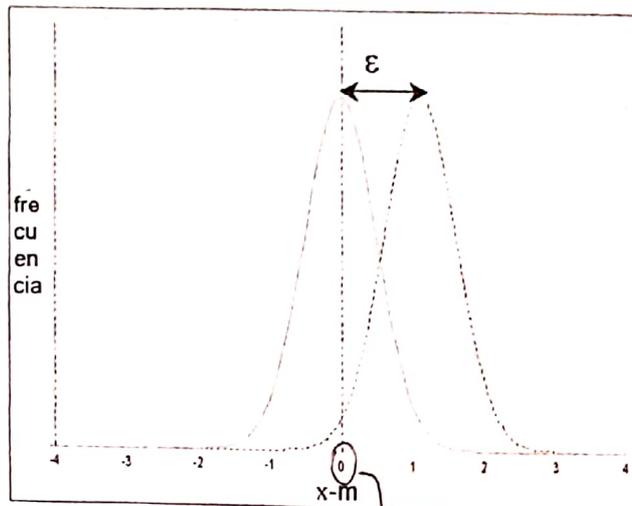
La evidencia experimental indica que los *errores aleatorios* tienen dos características importantes:

- i. tienen la misma probabilidad de aparecer resultados mayores que resultados menores a la media poblacional; esto es, la media es el valor más probable
- ii. el error aleatorio causará más a menudo desviaciones pequeñas que desviaciones grandes; es decir, es muy poco probable que el error aleatorio sea responsable de que un resultado del análisis se desvíe significativamente del valor medio.

Supongamos que se determina el contenido de hierro de un suplemento alimentario un número infinitamente grande de veces ( $n \rightarrow \infty$ ), siempre en las mismas condiciones; se calcula la media poblacional,  $\mu$ , y las desviaciones de cada valor respecto de la media poblacional,  $(x_i - \mu)$ . El verdadero valor de contenido de hierro seguiremos denominándolo  $\mu_0$ . Si graficamos la frecuencia (número de

veces que se repite una dada desviación) para cada una de las desviaciones observadas, obtendríamos una curva de Gauss o curva normal de error:

*nº veces que se repite una desviación*

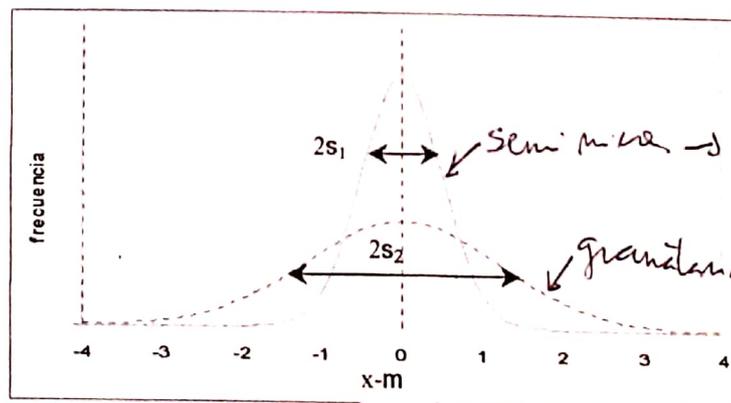


*← de variaciones observadas  
x̄: valor más probable*

La curva de línea llena describe la distribución de los errores aleatorios para un conjunto de valores exentos de error sistemático (donde  $\mu = \mu_0$ ). Las desviaciones respecto de la media se distribuyen simétricamente alrededor del cero. Las medidas que tienen desviación cero son las que se presentan con mayor frecuencia.

Supongamos ahora que la pipeta empleada para medir el volumen de muestra no fue previamente calibrada y el volumen real era mayor que el indicado. En consecuencia todas las determinaciones arrojarán un valor individual de hierro mayor que el verdadero. El efecto de éste error sistemático ( $> 0$ ) uniforme no descubierto sobre la curva normal se representó por la línea punteada. Mientras que el perfil de la curva no es alterado, la presencia del error desplazó a todos los valores hacia valores más positivos, en consecuencia la media estará desplazada y alejada del verdadero valor, es decir,  $\mu > \mu_0$ . La magnitud de este desplazamiento indica la magnitud del error sistemático.

Para ilustrar las propiedades de la curva normal de error veamos otro ejemplo. Supongamos que se mide la masa de un sólido con una balanza granataria (que pesa a  $\pm 0,01g$ )  $n$  veces ( $n \rightarrow \infty$ ), y luego con una balanza semi-micro (que pesa a  $\pm 0,0001g$ ) también  $n$  veces; para ambas series de datos obtenemos las medias y los desvíos individuales. Cada serie de valores estará afectada por un error aleatorio, pero su magnitud será ciertamente menor para la serie medida con balanza semi-micro debido a la mayor precisión que se consigue con este instrumento. El efecto se manifestará en el aspecto de las respectivas curvas de error. La figura que sigue refleja esto.



*(Error aleatorio)  
cantidad de desviaciones grandes respecto a la media*

El método de pesada usando la balanza semimicro dará una curva más alta y estrecha como resultado de un error aleatorio más pequeño que el método que empleó una balanza granataria. No obstante la diferencia en la forma, ambas curvas son simétricas alrededor del verdadero valor y las desviaciones grandes respecto de la media son mucho menos probables que las pequeñas desviaciones. La diferencia entre ambas series se refleja en el ancho de la curva.

La ecuación matemática que describe la curva normal de error es:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x_i-\mu)^2/2\sigma^2} \quad [8]$$

donde  $\mu$  es la media de un infinito número de mediciones,  $x_i - \mu$  es la desviación respecto de la media,  $f$  es la frecuencia con la que tiene lugar una dada desviación; y  $\sigma$  es la desviación estándar de la población, la cual tiene un único valor para una serie de datos, y determina el ancho de la curva normal de error. De hecho, el ancho de la campana en el punto de inflexión es igual a  $2\sigma$ . Así, el valor de  $\sigma$  es una medida muy útil de la desviación de un conjunto de datos. Para un número grande de datos se lo determina como

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n}} \quad [9]$$

Cualquiera sean los valores de  $\mu$  y  $\sigma$  de una serie de datos, se puede demostrar que el 95,5% de los valores caerán entre  $x_i = \mu + 2\sigma$  y  $x_i = \mu - 2\sigma$  y el 99,7% de los datos caerán entre  $x_i = \mu + 3\sigma$  y  $x_i = \mu - 3\sigma$

Tratamiento de un número pequeño de datos. Límites de confianza.

La curva normal nunca se obtiene experimentalmente porque demandaría un número enorme de mediciones. En la práctica real, en el laboratorio se realizan usualmente entre 2 y 6 mediciones para un análisis. En este caso, se calcula la media como el mejor valor para ese conjunto de mediciones y la precisión se calcula de

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} \quad [10]$$

La cantidad  $s$  es un "estimador de la desviación estándar de la población,  $\sigma$ ", para una serie finita de datos experimentales. Esto es,  $s \rightarrow \sigma$ , cuando  $n \rightarrow \infty$ . Supongamos el ejemplo siguiente. Se practicaron cinco determinaciones del contenido de CaO en una muestra de calcita,  $\text{CaCO}_3$ . Los valores individuales son:

Experimento	%CaO	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
1	55,95	0,11	0,0121
2	56,00	0,06	0,0036
3	56,04	0,02	0,0004
4	56,23	0,17	0,0289
5	56,08	0,02	0,0004
$\bar{x}$	56,06		$\Sigma = 0,0454$

Para este conjunto de cinco datos,  $\bar{x} = 56,06$  y  $s = (0,0454/4)^{1/2} = 0,11$ . Estos valores de  $\bar{x}$  y  $s$ , sin embargo pueden diferir sensiblemente de  $\mu$  y  $\sigma$ , respectivamente. Puede demostrarse que la mejor estimación de  $\mu$  está dada por el intervalo de confianza de la media:

$$\mu = \bar{x} \pm t \frac{s}{\sqrt{n}} \quad [11]$$

donde  $\bar{x}$ ,  $s$  y  $n$  conservan su significado; y  $t$  es la "t de Student", un parámetro obtenido de tabla y que depende del número de grados de libertad,  $\phi$  ( $\phi = n-1$ ) y del grado de confianza que se desee para el

resultado. El término  $t \times \frac{s}{\sqrt{n}}$  de la ecuación [11] define los "límites de confianza de la media". En nuestro ejemplo del %CaO, en que se practicaron 5 ensayos ( $\phi = 4$ ), se puede buscar en la tabla de t de Student, el valor de t.

Valores de t para varios niveles de probabilidad					
Grados de libertad	Factor para el intervalo de confianza				
	80%	90%	95%	99%	99,9%
1	3,08	6,31	12,7	63,7	637
2	1,89	2,82	4,30	9,92	31,6
3	1,64	2,35	3,18	5,84	12,9
4	1,53	2,13	2,78	4,60	8,60
5	1,48	2,02	2,57	4,03	6,86
6	1,44	1,94	2,45	3,71	5,96
7	1,42	1,90	2,36	3,50	5,40
8	1,40	1,86	2,31	3,36	5,04
9	1,38	1,83	2,26	3,25	4,78
10	1,37	1,81	2,23	3,17	4,59
11	1,36	1,80	2,20	3,11	4,44
12	1,36	1,78	2,18	3,06	4,32
13	1,35	1,77	2,16	3,01	4,22
14	1,34	1,76	2,14	2,98	4,14
∞	1,29	1,64	1,96	2,58	3,29

Por lo tanto:

con un nivel de confianza de 90%,  $\mu = 56,06 \pm 2,13 \frac{0,11}{\sqrt{5}} = 56,06 \pm 0,11$

*> confianza  
> límites de confianza*

con un nivel de confianza de 95%,  $\mu = 56,06 \pm 2,78 \frac{0,11}{\sqrt{5}} = 56,06 \pm 0,14$

y con un nivel de confianza de 99%,  $\mu = 56,06 \pm 4,60 \frac{0,11}{\sqrt{5}} = 56,06 \pm 0,23$

*> he mediciones  
t <*

Al aumentar la confianza de que el valor real esté comprendido en el intervalo definido, los límites de confianza se hacen mayores. Al aumentar el número de mediciones, no sólo se divide por un número mayor sino que además los valores de t de Student se reducen.

### 6. Propagación de errores aleatorios

En cualquier experimento se debe predecir la reproducibilidad que se puede esperar en el resultado final que se ha calculado a partir de dos o más medidas independientes, cada una de las cuales tiene asociada un dado error aleatorio. La forma en la que estos errores aleatorios individuales afecta al resultado final depende del tipo de cálculo.

#### 1) Adición y sustracción

En adiciones y/o sustracciones, la incertidumbre esperada en el resultado se obtiene de la raíz cuadrada de los cuadrados de los **errores absolutos** asociados a cada lectura. Ejemplo:

$$\begin{array}{r} 1,76 (\pm 0,03) \leftarrow e_1 \\ + 1,89 (\pm 0,02) \leftarrow e_2 \\ -0,59 (\pm 0,02) \leftarrow e_3 \\ \hline \end{array}$$

$$3,06 (\pm ?)$$

El resultado aritmético de  $1,76+1,89-0,59 = 3,06$ . Suponiendo que los errores absolutos en las mediciones son las indicadas entre paréntesis en cada caso, la incertidumbre asociada al resultado ( $e_4$ ) será:

$$e_4 = \sqrt{e_1^2 + e_2^2 + e_3^2} \quad [12]$$

en este caso:

$$e_4 = \sqrt{(0,03)^2 + (0,02)^2 + (0,02)^2} = \pm 0,04$$

La incertidumbre absoluta asociada al resultado de la suma algebraica es  $\pm 0,04$ , y el resultado puede expresarse como  $3,06 \pm 0,04$ .

2) **Multiplicación y División**

Cuando en el resultado final intervienen multiplicaciones y/o divisiones, se propagan los errores **relativos**. La incertidumbre en el resultado se calcula como sigue:

$$er_4 = \sqrt{er_1^2 + er_2^2 + er_3^2} \quad [13]$$

Ejemplo:

$$\frac{1,54(\pm 0,02) \times 3,82(\pm 0,04)}{1,23(\pm 0,03)} = 4,78 (\pm ??)$$

Los errores absolutos (indicados entre paréntesis) deberán primero transformarse en relativos ( $er_i$ ), luego se estima  $er_4$  según la ecuación [13].

$$er_1 = 0,02/1,54 = 0,0130$$

$$er_2 = 0,04/3,82 = 0,0105$$

$$er_3 = 0,03/1,23 = 0,0244$$

$$er_4 = \sqrt{(0,0130)^2 + (0,0105)^2 + (0,0244)^2} = \pm 0,0296$$

La incertidumbre absoluta en el dato calculado será  $= 0,0296 \times 4,78 = 0,141$ . El resultado final se expresa como  $4,8 \pm 0,1$ . Notar que no es conveniente redondear el resultado hasta que los cálculos se hayan terminado. Recién en el resultado final deberán considerarse las cifras significativas, es decir, conservar sólo los dígitos que sean significativos en el resultado obtenido.

3) **Operaciones combinadas**

Como último ejemplo consideremos la siguiente combinación de operaciones:

$$\frac{[1,76(\pm 0,03) - 0,59(\pm 0,02)]}{1,89(\pm 0,02)} = 0,619 \pm ??$$

Se deberá evaluar primero la diferencia en el numerador, utilizando las incertidumbres absolutas:

$$1,76 (\pm 0,03) - 0,59 (\pm 0,02) = 1,17 \pm 0,036$$

dado que  $\sqrt{(0,03)^2 + (0,02)^2} = 0,036$

Entonces se obtienen las incertidumbres relativas:

$$er_1 = 0,036/1,17 = 0,031$$

$$er_2 = 0,02/1,89 = 0,011$$

y finalmente,

$$er_3 = \sqrt{(0,031)^2 + (0,011)^2} = 0,033$$

la incertidumbre relativa en el resultado es 0,033 La incertidumbre absoluta será  $0,033 \times 0,619 = 0,02$ . El resultado debe expresarse como:  $0,62 \pm 0,02$ .

7. **Rechazo de resultados "anómalos"**

Cuando una serie de resultados obtenidos de n experimentos contiene un dato que resulta sospechoso por diferir excesivamente de la media, se debe decidir si retener o rechazar ese dato. La elección de un criterio de rechazo de un resultado sospechoso tiene riesgos. Se puede seguir un criterio muy estricto, rechazando datos que pueden pertenecer a la serie de resultados, o viceversa. Desafortunadamente no se puede invocar una regla universal para resolver la cuestión con un único criterio.

Se han desarrollado varios tests para dar criterios de rechazo o retención de datos sospechosos. Uno de ellos es el test Q de Dixon, que si bien puede resultar sencillo de aplicar, se debe tener presente que dado que se aplica a muestras que usualmente contienen pocos datos experimentales, se deben usar sólo como un elemento de ayuda al sentido común.

Para aplicar el test Q a un dato sospechoso, primero los datos se ordenan de menor a mayor, luego se calcula la diferencia entre el valor cuestionable ( $x_q$ ) y su vecino más próximo ( $x_p$ ), y finalmente, se calcula la diferencia entre los valores máximo y mínimo de la serie o rango de datos.

$$Q_{\text{experimental}} = \frac{|x_q - x_n|}{\text{rango}} \quad [14]$$

Este valor de  $Q_{\text{exper}}$  se compara con los valores de rechazo  $Q_{\text{critico}}$  que se informan en la tabla siguiente:

Nº de observaciones	Q <sub>crítico</sub>		
	90% de confianza	95% de confianza	99% de confianza
3	0,941	0,970	0,994
4	0,765	0,829	0,926
5	0,642	0,710	0,821
6	0,560	0,625	0,741
7	0,507	0,568	0,680
8	0,468	0,526	0,634
9	0,437	0,493	0,598
10	0,412	0,466	0,568

Si  $Q_{\text{exper}}$  es mayor que  $Q_{\text{crítico}}$  el resultado sospechoso puede ser rechazado con el grado de confianza indicado. Retomemos el ejemplo del análisis de %CaO. En la serie de datos, el valor 56,23 resulta bastante mayor que la media. Si aplicásemos el test Q a este dato tendríamos:

$$Q_{\text{experimental}} = \frac{|56,23 - 56,08|}{56,23 - 55,95} = \frac{0,15}{0,28} = 0,54$$

Para 5 observaciones, y un nivel de confianza de 95%, el  $Q_{\text{crítico}}$  es 0,71, mayor que el  $Q_{\text{experimental}}$ . El resultado debe conservarse en la serie. Observar que en éste ejemplo existe un 5% de probabilidad de conservar el dato siendo anómalo.

Se pueden sugerir algunas recomendaciones para el tratamiento de una serie reducida de datos que contienen un valor sospechoso:

1. reexaminar cuidadosamente todas las observaciones incluidas el dato sospechoso para asegurar que un error grosero no sea la causa de ese dato anómalo
2. estimar la precisión esperada del procedimiento para asegurarse que el resultado discrepante realmente es cuestionable
3. repetir los análisis si se dispone de muestra suficiente. Si la nueva serie coincide con la anterior aumentará el peso de la sospecha de que el resultado anómalo debería rechazarse. Por otra parte si el criterio de rechazo aún no se cumple y el dato debe retenerse, su incidencia en la media global se minimiza

En síntesis, la única razón válida para rechazar un dato de una serie reducida de datos es el convencimiento de que se ha cometido un error en el proceso de medida. Sin este convencimiento, lo más prudente es mantener una actitud cautelosa y conservar al dato sospechoso.

## 8. Errores sistemáticos y aleatorios en el análisis volumétrico

Se puede considerar que un análisis volumétrico completo involucra etapas de pesada de muestra sólida y mediciones de volumen con pipeta aforada y con bureta. Examinaremos los errores aleatorios y sistemáticos que pueden cometerse con el uso de una balanza y aquellas que implican el uso de material volumétrico.

### Errores asociados a una pesada.

Los errores aleatorios asociados al procedimiento de pesada son, en general, pequeños. En el caso de una balanza semi-micro que pesa a la décima de mg ( $\pm 0,1$  mg), la incertidumbre de cada lectura es de  $\pm 0,1$  mg. Debido a que, regularmente, la cantidad que se pesa se encuentra en alrededor de 1 g o más, resulta evidente que el error aleatorio relativo a la pesada efectuada será bajo:  $(\pm 0,0001/1) \cdot 100 = 0,01\%$ . Este error será mayor cuando se reduce la masa a pesar.

Los errores sistemáticos en una pesada pueden ser apreciables y proceder de diversas fuentes. Los más importantes son:

- i. adsorción de humedad sobre la superficie del recipiente de pesada
- ii. el no permitir que el objeto a pesar alcance la misma temperatura que la balanza antes de pesar (error muy frecuente en gravimetría, cuando se pesan crisoles)
- iii. pesas oxidadas o contaminadas con polvo
- iv. efecto del empuje del aire que actúa en forma diferente en el platillo de las pesas y en el del objeto a pesar

Con el fin de lograr mayor exactitud en las pesadas, balanzas y pesas del laboratorio deben ser calibradas por autoridades de control. Aparte de usar instrumentos calibrados, algunas

precauciones experimentales sencillas para minimizar errores sistemáticos durante una pesada incluyen: 1) estabilizar la temperatura del objeto a pesar con la del laboratorio; 2) realizar pesadas por diferencia, es decir, se pesa primero un pesafiltro conteniendo la muestra a pesar y luego sin ella. La diferencia de peso entre las dos pesadas proporciona el peso de la muestra. Este procedimiento evita errores sistemáticos como por ejemplo, humedad o polvo en la superficie del pesafiltro (el que deberá manipularse evitando tocar con las manos directamente).

#### Errores asociados al material volumétrico.

Los errores *aleatorios* en volumetría provienen de la incertidumbre asociada al uso de material de vidrio. Al llenar un matraz aforado de 250 mL hasta el aforo, existirá una incertidumbre en la lectura de volumen (es decir la distancia entre el menisco y la marca) que depende del diámetro interno del cuello del matraz. Lo mismo ocurre con pipetas y buretas. La Tabla siguiente indica los errores aleatorios máximos que se asocian a algunos de los materiales que usualmente se emplearán en este laboratorio.

Material volumétrico	Error absoluto en c/lectura	Error absoluto del material
Matraz (10 mL)	±0,02 mL	±0,02 mL
Matraz (50 mL)	±0,05 mL	±0,05 mL
Matraz (100 mL)	±0,08 mL	±0,08 mL
Matraz (250 mL)	±0,12 mL	±0,12 mL
Matraz (500 mL)	±0,20 mL	±0,20 mL
Matraz (1 L)	±0,30 mL	±0,30 mL
Pipetas doble aforo (10 mL)	±0,02 mL	±0,03 mL
Pipetas doble aforo (25 mL)	±0,03 mL	±0,04 mL
Buretas	±0,02 mL	±0,03 mL
Balanza analítica (a 0,1 mg)	±0,1 mg	±0,14 mg

La segunda columna de la Tabla indica el error aleatorio máximo en la lectura de volumen o masa de algunos instrumentos usuales en el trabajo analítico. La tercer columna indica el error aleatorio asociado al material. Como en el manejo de buretas y pipetas de doble aforo son dos las lecturas a efectuar, la incertidumbre debe calcularse según la regla de propagación de sumas y restas. Es decir, en el caso de una bureta, existe incertidumbre en el enrase del menisco al comienzo de la titulación (±0,02 mL) y en la lectura final de volumen, por lo que la incertidumbre asociada a la bureta será:

$$e_{\text{bureta}} = \sqrt{(\pm 0,02)^2 + (\pm 0,02)^2} = \pm 0,028 \approx \pm 0,03 \text{ mL}$$

De igual modo, para una balanza analítica de sensibilidad de ±0,1 mg, la incertidumbre absoluta asociada a la masa pesada será:

$$e_{\text{pesada}} = \sqrt{(\pm 0,1)^2 + (\pm 0,1)^2} = \pm 0,14 \text{ mg}$$

Los *errores sistemáticos* involucrados en procedimientos volumétricos pueden tener varios orígenes. Por citar a los más importantes:

- i. error en el vaciado de pipetas y buretas, es decir error por operación incorrecta de estos materiales
- ii. error en la calibración del material de vidrio. Por ejemplo suponer que un matraz es de 250,00 mL y en realidad el aforo corresponde a 249,90 mL. En este caso, el error en el volumen se reflejará en el resultado final de todos los experimentos basados en el uso de dicho matraz.
- iii. error asociado a diferencia en la temperatura de calibración del material de vidrio y la del experimento
- iv. finalmente un error importante en el análisis volumétrico es el error del "indicador". Este error puede ser importante en magnitud, comparado con los errores aleatorios. Este error, que se estudiará en detalle en cada una de las titulaciones que se realicen, consiste en el exceso de volumen de titulante que se debe adicionar para observar un cambio de color de la solución. Rápidamente, si se valora un ácido fuerte empleando una base fuerte, el pH en el punto de equivalencia será neutro (pH=7). Si se emplea fenolftaleína como indicador de punto final, el

viraje de la fenolftaleína es en el rango de pH de 8,0-9,6, es decir, se deberá agregar un pequeño exceso de base para observar un viraje de color en la solución.

Se deduce que: a) a diferencia de los errores aleatorios, los sistemáticos no se pueden apreciar con la simple repetición de mediciones y b) a menos que se conozca anticipadamente el valor más probable del resultado de un análisis (situación sumamente improbable), pueden existir errores sistemáticos muy grandes que pasen inadvertidos si no se toman las debidas precauciones. En resumen, los errores sistemáticos asociados a un experimento pueden evitarse y/o corregirse. Para esto, se pueden usar muestras patrones, realizar mediciones con un método independiente, realizar ensayos en blanco, cambiar el tamaño de la muestra.

Los errores aleatorios no pueden corregirse. Existen, sin embargo, un conjunto de métodos estadísticos que permiten estimarlos, siguiendo las reglas de propagación de errores aleatorios, para lo cual se supone que los errores sistemáticos fueron detectados y corregidos con anterioridad. El siguiente cálculo nos permite predecir la incertidumbre en la molaridad de una solución valorada contra una droga patrón primario. La molaridad de una solución se calcula a partir de una serie de mediciones que incluyen determinación de la masa del patrón primario por pesada y luego lectura de volumen de titulante gastado hasta alcanzar el punto de equivalencia de la titulación. La molaridad se calcula como:

$$M = \frac{w (\pm e_{\text{pesada}})}{pm \times V (\pm e_{\text{bureta}})} \quad [15]$$

w: masa de patrón primario pesada en gramos con un error aleatorio de  $\pm 0,00014$  g

pm: peso molecular del patrón en g/mol. En el análisis volumétrico se considera que los pesos moleculares tiene un error despreciable en relación con los asociados a la pesada y a la lectura de volumen.

V: volumen de titulante gastado (en mL), con un error de  $\pm 0,03$  mL

La incertidumbre en la molaridad se podrá estimar como:

$$e_r(\text{Molaridad}) = \sqrt{(e_{\text{pesada}}/w)^2 + (e_{\text{bureta}}/V)^2}$$

$$e(\text{Molaridad}) = M \times e_r(\text{Molaridad})$$

Ejemplo: Se valoró una solución de NaOH empleando biftalato de potasio como patrón primario (peso molecular=204,22 mg/mmol). Para ello se pesaron 432,7 ( $\pm 0,1$ ) mg de droga sólida previamente secada en estufa, se disolvió el sólido en agua destilada y se valoró con la solución de NaOH gastándose 19,35 mL. La molaridad será:

$$M(\pm ??) = \frac{432,7(\pm 0,1)}{204,22 \times 19,35(\pm 0,03)} = 0,1094984 \text{ M}$$

y la incertidumbre en el resultado:

$$e_r(\text{Molaridad}) = \sqrt{(0,1/432,7)^2 + (0,03/19,35)^2} = 1,57 \times 10^{-3}$$

$$e(\text{Molaridad}) = 0,1094984 \times 1,57 \times 10^{-3} = 0,00017 \approx 0,0002$$

Por lo que la molaridad del NaOH se debe expresar como: (0,1095  $\pm$  0,0002) M.

## 9. Convenio sobre el uso de cifras significativas

El ejemplo de valoración de NaOH demuestra como se deben expresar los resultados analíticos de una valoración. El cálculo del error asociado a la molaridad del NaOH es de 0,00017, que debe aproximarse a 0,0002, dado que el cuarto lugar decimal ya es dudoso. A su vez, el cálculo de la molaridad deberá acotarse en sus cifras significativas: si el cuarto decimal es el dudoso, la molaridad deberá expresarse en este ejemplo con cuatro cifras significativas: el resto de los dígitos obtenidos de la calculadora deben aproximarse al último dígito significativo, en este caso se redondea a 0,1095 M. Como regla general, un resultado obtenido con todos los dígitos de la calculadora que debe luego acotarse de acuerdo con el error aleatorio asociado. Pero esto se hará recién cuando se escriba el resultado final.

# 4F PREGUNTAS Y PROBLEMAS

- 4-1. Explíquese la diferencia entre:
- \*a) exactitud y precisión
  - b) error aleatorio y error sistemático
  - \*c) media y mediana
  - d) error absoluto y error relativo
  - \*e) varianza y desviación estándar

- 4-2. Defínase:
- \*a) rango
  - b) coeficiente de variación
  - \*c) cifras significativas
  - d) distribución gaussiana

Para cada serie calcúlese:

- a) la media
- b) la mediana
- c) la dispersión o rango
- d) la desviación estándar
- e) el coeficiente de variación
- f) la desviación estándar estimada de la dispersión

- 4-5. Los valores aceptados para la serie de datos del problema 4-4 son: \*serie A, 2.0; serie B, 69.75; \*serie C, 0.0930; serie D, 3.0; \*serie E, 69.05; serie F, 0.635. Para cada serie calcúlese:

- a) el error absoluto
- b) el error relativo en partes por mil.

- \*4-6. Estímese la desviación estándar absoluta y el coeficiente de variación para los resultados de los siguientes cálculos. Redondéese cada resultado de modo que sólo contenga cifras significativas. Los números entre paréntesis son las desviaciones estándar absolutas.

a)  $y = 6.75(\pm 0.03) + 0.843(\pm 0.001) - 7.021(\pm 0.001)$   
 $y = 0.572$

b)  $y = 19.97(\pm 0.04) + 0.0030(\pm 0.0001) +$   
 $y = 1.29(\pm 0.08) = 21.263$

c)  $y = 67.1(\pm 0.3) \times 1.03(\pm 0.02) \times 10^{-17}$   
 $y = 6.9113 \times 10^{-16}$

d)  $y = 243(\pm 1) \times \frac{760(\pm 2)}{1.006(\pm 0.006)} = 183,578.5$

e)  $y = \frac{143(\pm 6) - 64(\pm 3)}{1249(\pm 1) + 77(\pm 8)} = 5.9578 \times 10^{-2}$

f)  $y = \frac{1.97(\pm 0.01)}{243(\pm 3)} = 8.106996 \times 10^{-3}$

- 4-7. Estímese la desviación estándar absoluta y el coeficiente de variación para los resultados de los siguientes cálculos. Redondéese cada resultado de manera que sólo se incluyan cifras significativas. Los números entre paréntesis son las desviaciones estándar absolutas.

a)  $y = -1.02(\pm 0.02) \times 10^{-7} - 3.54(\pm 0.2) \times 10^{-4}$   
 $= -1.374 \times 10^{-7}$

b)  $y = 100.20(\pm 0.08) - 99.62(\pm 0.06) +$   
 $0.200(\pm 0.004) = 0.780$

c)  $y = 0.0010(\pm 0.0005) \times 18.10(\pm 0.02) \times 200(\pm 1)$   
 $= 3.62$

d)  $y = \frac{1.73(\pm 0.03) \times 10^{-14}}{1.63(\pm 0.04) \times 10^{-16}} = 106.1349693$

e)  $y = \frac{100(\pm 1)}{2(\pm 1)} = 50$

f)  $y = \frac{1.43(\pm 0.02) \times 10^{-2} - 4.76(\pm 0.06) \times 10^{-3}}{24.3(\pm 0.7) + 8.06(\pm 0.08)}$   
 $= 2.948 \times 10^{-4}$

- 4-8. Redondéese cada uno de los siguientes resultados y que sólo incluyan cifras significativas.

\*a)  $y = \log 1.73 = 0.238046$

b)  $y = \log 0.0432 = -1.364516$

\*c)  $y = \log (6.022 \times 10^{23}) = 23.77960$

d)  $y = \log (4.213 \times 10^{-21}) = -20.375409$

\*e)  $y = \text{antilog} (-3.47) = 3.38844 \times 10^{-4}$

- \*4-3. Sugieranse algunas fuentes de error aleatorio al medir el ancho de una tabla de 3 m con una regla metálica de 1 m.

- 4-4. Considérense las siguientes series de mediciones repetidas:

*A	B	*C	D	*E	F
2.4	69.94	0.0902	2.3	69.65	0.624
2.1	69.92	0.0884	2.6	69.63	0.613
2.1	69.80	0.0886	2.2	69.64	0.596
2.3		0.1000	2.4	69.21	0.607
1.5			2.9		0.582

f)  $y = \text{antilog } 5.7 = 5.01187 \times 10^5$

\*g)  $y = \text{antilog } 0.99 = 9.77237$

h)  $y = \text{antilog} (-27.2424) = 5.722687 \times 10^{-28}$

- \*4-9. El análisis para ion potasio de varias preparaciones para alimentos vegetales dio los siguientes datos:

Muestra	Porcentaje medio de K <sup>+</sup>	Número de observaciones	Desviación de los resultados individuales respecto de la media
1	4.80	5	0.13, 0.09, 0.07, 0.05, 0.06
2	8.04	3	0.09, 0.08, 0.12
3	3.77	4	0.02, 0.15, 0.07, 0.10
4	4.07	4	0.12, 0.06, 0.05, 0.11
5	6.84	5	0.06, 0.07, 0.13, 0.10, 0.09

- a) Evalúese la desviación estándar  $s$  para cada muestra
- b) obtenga un estimado ponderado para  $s$

- 4-10. Se analizaron seis botellas de vino para azúcar residual y se obtuvieron los siguientes resultados:

Frasco	Porcentaje (p/v) de azúcar residual	Número de observaciones	Desviación de los resultados individuales respecto de la media
1	0.94	3	0.050, 0.10, 0.08
2	1.08	4	0.060, 0.050, 0.090, 0.060
3	1.20	5	0.05, 0.12, 0.07, 0.00, 0.08
4	0.67	4	0.05, 0.10, 0.06, 0.09
5	0.83	3	0.07, 0.09, 0.10
6	0.76	4	0.06, 0.12, 0.04, 0.03

- a) Evalúese la desviación estándar  $s$  para cada serie de datos
- b) Agrupe los datos para determinar una desviación estándar absoluta para el método

- \*4-11. Se analizaron por duplicado nueve muestras de heroína ilegal por medio de cromatografía de gases. Agrupe los siguientes datos para determinar una desviación estándar absoluta para el procedimiento:

Muestra	Heroína (%)	Muestra	Heroína (%)
1	2.24, 2.27	6	1.07, 1.02
2	8.4, 8.7	7	14.4, 14.8
3	7.6, 7.5	8	21.9, 21.1
4	11.9, 12.6	9	8.8, 8.4
5	4.3, 4.2		

4-12. Calcúlese la  $s$  ponderada para el siguiente análisis espectro-fotométrico para ANT (ácido nitrilotriacético) en agua del río Ohio:

Muestra	ANT, ppb
1	13, 16, 14, 9
2	38, 37, 38,
3	25, 29, 23, 29, 26

4-13. El cambio de color de un indicador químico necesita una sobretitulación de 0.03 mL. Calcúlese el porcentaje de error relativo si el volumen total de titulante es:

- \*a) 50.00 mL
- b) 10.0 mL
- \*c) 25.0 mL
- d) 40.0 mL

4-14. En el curso de un análisis para Zn hay una pérdida de 0.4 mg del elemento. Calcúlese el porcentaje de error relativo debido a esta pérdida si el peso de Zn en la muestra es:

- \*a) 40 mg
- b) 175 mg
- \*c) 400 mg
- \*d) 600 mg