TRATAMIENTO DE LOS RESULTADOS ANALÍTICOS

Aplicación de la estadística en el laboratorio

Respuestas de los ejercicios de evaluación

El objetivo de este documento es ayudar al profesorado y al lector en general, proporcionando las respuestas a las unidades de evaluación que contiene el libro *TRATAMIENTO DE LOS RESULTADOS ANALÍTICOS. Aplicación de la estadística en el laboratorio* de Joan Sánchez y Miquel Villalobos con ISBN 978-8496960-48-0.

ACTIVIDAD 8
UNIDAD DE EVALUACIÓN

EJERCICIOS DE EVALUACIÓN

8.1

1) *V*

Se cumple la propiedad distributiva: $A \cdot (B \pm c) = A \cdot B \pm A \cdot c$
Para comprobarlo trataremos la operación como un producto y seguiremos las reglas correspondientes:
$5 \cdot (1{,}02 \pm 0{,}01) = R_F \pm I_{RF}$
$R_F = 5 \cdot 1{,}02 = 5{,}10$

$$\frac{I_{RF}}{R_F} = \sqrt{(0/5)^2 + (0{,}01/1{,}02)^2} = \frac{0{,}01}{1{,}02}$$

$$I_{RF} = \frac{0{,}01}{1{,}02}(1{,}02 \cdot 5) = 0{,}05$$

2) *F*

Como máximo la medida que puede proporcionar es 200,00 g que contiene cinco cifras significativas.

3) *F*

Como máximo, el valor que puede proporcionar la medida es 100,000 g que contiene seis cifras significativas.

4) *F*

Existen diversas incorrecciones. En primer lugar 1,05 contiene dos decimales pero tres cifras significativas. Además, si redondeamos a dos cifras significativas, el resultado corresponde a 1,0

5) *V*

Redondeando a dos cifras significativas (y por tanto, a un decimal) se nos presenta la duda de expresar el resultado como 1,0 o 1,1. De acuerdo al criterio del valor par, el resultado es 1,0

6) *F*

Para una multiplicación, el resultado final ha de contener el número de cifras significativas correspondientes a las del factor que presenta menos, en este caso, 10

7) *V*

Se ha de redondear una sola vez, después se ha de realizar todos los cálculos, para no introducir errores por truncamiento.

8) *V*

En un instrumento es habitual medir la fiabilidad a partir del número de horas de funcionamiento en que se comporta correctamente, es decir, con los grados de precisión y exactitud previstos.

9) *F*

No prestar atención en las operaciones de enrase a partir de la posición del menisco de los líquidos constituye, ciertamente, una fuente de error. En este caso al no ser conscientes del procedimiento aplicado el error se puede efectuar por exceso o por defecto, sin una tendencia determinada y por este motivo clasificarlo como error aleatorio.

10) ***F***
La media aritmética es el parámetro de centralización más utilizado y muestra el valor de la tendencia central de la serie de valores. No determina el grado de dispersión de los datos.

11) ***V***
El grado de exactitud se determina a partir de la distancia existente (error) entre el representante de la serie de datos experimentales y el valor considerado como cierto. Este error se puede expresar matemáticamente de forma absoluta o relativa pero, en cualquier caso, al aumentar su valor disminuye el grado de exactitud.

12) ***V***
Contra más resolución presente un instrumento, más precisión mostrará en las mediciones que se realicen con él. Dicho de otra forma, al aumentar la resolución, aumenta el número de cifras significativas que proporciona y por tanto su precisión.

13) ***F***
La exactitud y la precisión son propiedades independientes que presenta una serie de mediciones. El grado de exactitud no depende del grado de precisión. Un método exacto puede resultar muy o poco preciso.

14) ***F***
Resulta válido el razonamiento anterior.

15) ***F***
La reproductibilidad es una medida de la precisión obtenida bajo condiciones diferentes. En la obtención de los valores de las medidas no coinciden los laboratorios, los instrumentos, los técnicos, etc.

16) ***V***
La fiabilidad de un instrumento o aparato aumenta con el número de horas de trabajo en que opera de forma satisfactoria, manteniendo el grado de exactitud y precisión esperadas.

17) ***V***
La mayor parte de los animales de experimentación responden de igual forma a una misma dosis de un fármaco. Sólo una pequeña fracción de la población presentará menos efecto y otra pequeña fracción de la población presentará un efecto más intenso con la misma dosis suministrada. La distribución resultante del estudio del efecto frente a la dosis se puede considerar normal (en general). Basándonos en este comportamiento, se determinan las dosis efectivas de tratamiento de los fármacos, las dosis perjudiciales, la dosis letal de los tóxicos, etc.

18) ***F***
Un determinado error humano puede seguir una tendencia determinada (error sistemático) o, por el contrario, presentar un comportamiento inesperado (error aleatorio). Sólo los errores experimentales aleatorios, de origen humano o no, siguen una distribución normal.

19) ***F***
Sólo en el caso en que estos errores de método no sean sistemáticos seguirán una distribución normal.

20) ***F***
La calidad de un método analítico se evalúa a partir del estudio de toda una serie de parámetros analíticos y no analíticos. Las herramientas más utilizadas para comprobar los principales parámetros (exactitud, precisión, sensibilidad, límite de detección, etc.) suelen basarse en técnicas de regresión o en la aplicación de los correspondientes tests estadísticos.

21) *V*

La función normal incluye en su expresión un factor de normalización $(1/s\sqrt{2\pi})$ *para cumplir la condición citada, es decir, que el área que determina la curva sobre el eje X valga 1*

22) *F*

El valor medio está siempre situado en el centro de una curva normal.

23) *V*

La totalidad del área incluida debajo de la curva normalizada de Gauss vale 1 (o el 100 %).

24) *F*

En una distribución con asimetría positiva (en forma de L) se cumple que la moda es menor al valor medio. En relación a la mediana, su posición puede depender del grado de normalidad de la distribución.

25) *F*

El intervalo de confianza para una serie de mediciones está definido por:

$$\bar{x} \pm \frac{st}{\sqrt{N}}$$

Si aumenta N, disminuye el intervalo.

26) *V*

Sólo son aceptables los resultados que provienen de datos aceptables. No obstante se han de aplicar criterios objetivos de aceptación de resultados.

27) *V*

Una vez fijado el nivel de probabilidad (por ejemplo, del 95 %), al aumentar el número de datos aumenta el valor de R (se puede observar en las tablas).

28) *V*

Puesto que el intervalo de aceptación es más pequeño ($[\bar{x} \pm 2{,}5d] < [\bar{x} \pm 4d]$) *el criterio resulta más restrictivo y algunos valores que podrían ser rechazados con la aplicación del criterio 2,5 d podrían, en cambio, ser aceptados con el criterio 4d*

29) *F*

A pesar de ser uno de los criterios más extendidos dentro del campo de la química analítica para validar posibles valores anómalos, en el caso de pequeñas series de datos (que es el más habitual) tiende a ser poco estricto

30) *F*

Para una serie de datos, el cálculo de un intervalo de incertidumbre supone la definición de un determinado intervalo de confianza. Conceptualmente ambos intervalos se refieren a realidades diferentes: el intervalo de incertidumbre corresponde al valor de la expresión $st/\sqrt{N}$*, mientras el intervalo de confianza se define como* $\bar{x} \pm st/\sqrt{N}$

31) V

El número de grados de libertad se calcula como el número de datos menos el número de series. En este caso, n = N - 1

32) *F*

La expresión anterior muestra el intervalo de incertidumbre. El intervalo de confianza se determina con:

$$\bar{x} \pm \frac{st}{\sqrt{N}}$$

33) ***F***
El criterio de aceptación de Grubbs tiene una base probabilística y emplea tablas que dependen del nivel de confianza que exigimos.

34) ***V***
Los valores de Q (90%) disminuyen con el número de datos y son, por la propia definición del criterio, siempre inferiores a 1. Contrariamente, los valores de R (90 %) aumentan con el número de datos y son siempre superiores a la unidad.

35) ***F***
Para una serie de datos determinada, al aumentar el nivel de confianza aumenta el valor de t correspondiente y, por tanto, también el valor del intervalo. Además, la experiencia nos muestra que si aumenta el intervalo de confianza, debe aumentar la seguridad de que un determinado valor se encuentre incluido.

36) ***F***
Al aumentar la probabilidad se incrementa el intervalo de confianza y, en consecuencia, los límites de confianza inferior y superior se alejan del valor medio. La posición de los límites, sin embargo, depende del valor medio.

37) ***V***
Esta es una de las propiedades de las funciones normales. Una distribución teóricamente normal presenta dos puntos de inflexión, a la izquierda y derecha del valor medio, separados de este por una desviación estándar (± s).

38) ***V***
Independientemente del nivel de probabilidad elegido, R es siempre superior a 1

39) ***F***
El valor de Q (90 %) depende únicamente del número de datos pero el valor que toma no varía proporcionalmente con N

40) ***F***
La aplicación de los correspondientes tests estadísticos, por ejemplo, para evaluar posibles diferencias de precisión o exactitud, nada más confirman si las diferencias detectadas son o no significativas. En caso de que lo sean podremos afirmar, con el nivel de seguridad adoptado, que las muestras corresponden a poblaciones diferentes. En caso, sin embargo, que las diferencias no resulten significativas no podemos afirmar que las muestras pertenecen a la misma población.

41) ***F***
No siempre es así. En el caso de evaluar la posible diferencia de exactitud de dos series independientes, se aplica el correspondiente test de comparación empleando el estadístico t y los correspondientes valores medios. Si, en cambio, las dos series están aparejadas, es decir, si el valor que toma en una serie está condicionado por el valor que había tomado en la otra, la prueba estadística a aplicar es diferente.

42) ***V***
El test de xi cuadrado evalúa si las diferencias en cuanto a frecuencias entre dos series son o no significativas, a diferencia de los test t que evalúan diferencias de exactitud o del test F que evalúa diferencias de precisiones.

43) ***F***
El criterio de Fischer evalúa posibles diferencias de precisión entre dos series, no de exactitud.

44) ***F***
Por la propia definición de los estadísticos, tanto los valores de la F tabulada como de la F calculada son siempre superiores o iguales a 1

45) ***V***
La sensibilidad de calibración es única durante la aplicación de una determinada regresión y corresponde al valor de la pendiente de la función de ajuste, a diferencia de la sensibilidad analítica que depende del valor de cada patrón empleado.

46) ***F***
El valor del coeficiente r, en valor absoluto, varía teóricamente desde 0 (ningún grado de asociación entre las variables) a 1 (total asociación entre estas).

47) ***F***
A pesar de emplear los mismos datos en el ajuste, las hipótesis iniciales y el método de cálculo empleado para cada método hace que las ecuaciones resultantes sean diferentes. La elección de un método de ajuste u otro dependerá de los recursos disponibles y la utilización que se quiera realizar de la ecuación resultante.

48) ***F***
No siempre. De hecho, el método más extendido es el método alternativo, el método de los mínimos cuadrados, fácilmente aplicable con las calculadoras y hojas de cálculo habituales.

49) ***F***
Justamente la gran ventaja que incorpora el método de la mediana simple es su robustez delante de la presencia de valores anómalos.

50) ***V***
Debido al propio algoritmo en el tratamiento de los datos, el uso de la mediana simple hace que los valores anómalos vayan separándose de la tendencia general de los datos. Puede realizarse, por tanto, el estudio global de las datos a pesar de incorporar una porción de datos afectados de error.

51) ***F***
La precisión de una serie de datos es una evaluación de la aportación de los errores aleatorios que se han cometido. Matemáticamente, sin embargo, no se calcula como la suma directa de estos.

52) ***F***
En relación a la aplicación de un método analítico, la calibración con patrones se realiza determinando la mejor función existente entre los valores analíticos obtenidos y los valores de la concentración correspondientes a los patrones. La sensibilidad, entonces, se define como la medida de la capacidad que presenta la técnica por poder discriminar pequeñas diferencias de ***concentración*** *del analito.*

Fijémonos que si la pregunta fuera: "La sensibilidad es la capacidad de discriminar pequeñas diferencias cuantitativas de concentración de una determinada sustancia", la respuesta sería V

53) ***F***
Los CRM (Certified Reference Material) justamente son materiales estándares externos que se pueden adquirir para disponer de patrones de absoluta confianza en los procesos de validación de la calidad propios del laboratorio.

54) ***F***
Aunque a menudo se confunden estos dos conceptos, rigurosamente verificar supone realizar acciones internas de control en el laboratorio, mientras calibrar supone la intervención de técnicos especializados de una empresa homologada los cuales, finalmente, deben entregar el correspondiente certificado de conformidad.

55) ***F***
Ambos límites se sitúan en la misma zona de la curva de calibración, pero se refieren a conceptos diferentes. El límite de detección, por convenio, corresponde al valor de la concentración resultante del valor del señal analítico calculado como el señal correspondiente al blanco más tres veces su desviación estándar. El límite superior de la zona muerta representa, en cambio, el punto de transición entre la zona de respuesta no lineal a bajas concentraciones y la zona lineal en la que hay una proporcionalidad entre respuesta y concentración del analito.

56) ***F***
El límite de detección de la técnica está fuertemente determinado por el ruido sin embargo, en realidad, se trata de dos conceptos diferentes. El ruido es una señal extraña e indeseable que se superpone a la señal del analito a medir y supone la última limitación en relación a la exactitud y la sensibilidad que el método analítico puede proporcionar. El límite de detección, en cambio, corresponde a un determinado valor de concentración a partir del que podremos afirmar, con una cierta seguridad estadística, que la señal medida corresponde a la aportación del analito y no únicamente a la del ruido existente.

57) ***F***
Matemáticamente, las concentraciones correspondientes a los límites de cuantificación y detección se calculan como:
$C_{LQ} = 10\ s_B/m$ *y* $C_{LD} = 3\ s_B/m$*, siendo* s_B *la desviación estándar del blanco y* m *la pendiente de la recta de ajuste. Puede verificarse que:* $C_{LQ} = (10/3)\ C_{LD}$

58) ***F***
La zona de cuantificación dudosa corresponde a valores de concentraciones inferiores a $10s_B/m$*, mientras que la zona de detección aceptable corresponde a valores de concentración superiores a* $3s_B/10$*. Por debajo de valores de concentración* $3s_B/10$*, la zona debería ser descrita tanto de cuantificación como de detección dudosa.*

59) ***V***
Una parte en un trillón (ppt), tal como se define en términos de concentración, corresponde a la relación existente entre una parte y 10^{12}*, por ejemplo entre 1* μ*g y 1 kg*

60) ***F***
1 ppq corresponde a una parte entre 10^{15}

8.2

a)

Ventajas
El método proporciona una buena ecuación de ajuste a pesar de la presencia de algunos puntos anómalos. No es necesario, por tanto, efectuar una validación previa de los datos. El cálculo, aunque lento, no requiere de grandes recursos y puede efectuarse manualmente. El método permite la detección y selección de los puntos anómalos a partir del análisis de residuales. Este método puede constituir, en consecuencia, una primera etapa de

depuración de los valores experimentales si se decide, finalmente, emplear el método estándar de los mínimos cuadrados.

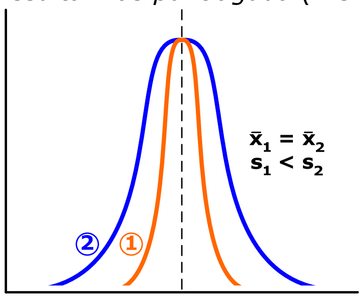
Inconvenientes
No es el método más empleado y su cálculo no está automatizado en calculadoras científicas y hojas de cálculo habituales. No se dispone de un buen parámetro para validar la calidad del ajuste. Con este método no se pueden obtener los intervalos de confianza correspondientes a los parámetros de ajuste de la curva de regresión, lo que limita la comparación entre dos técnicas analíticas por regresión.

b) *El límite de cuantificación de una técnica analítica se define como aquel valor mínimo de concentración del analito a partir del que es posible una cuantificación aceptable. Matemáticamente se calcula como:*
$C_{LQ} = 10\ s_B/m$
Donde s_B es la desviación estándar del blanco y m la pendiente de la recta de ajuste.

c) *Aunque existan más propiedades, resaltamos las siguientes:*
-El valor medio determina el centro de la distribución y la desviación estándar el grado de compactación de los datos alrededor de este.
- La curva normal es simétrica respeto a un eje vertical que corresponde, teóricamente, al valor medio, la moda y la mediana.
-La curva es asintótica con respecto al eje X el que describe que, para cualquier valor de la variable, siempre existirá una probabilidad diferente de cero de aparición del valor.
-El área incluida bajo la campana normal vale la unidad.

d) *La variancia es un parámetro de dispersión definido como el cuadrado de la desviación estándar:*
$V = s^2$.
Se utiliza habitualmente en el test de Fischer-Snedecor para evaluar si la diferencia de precisión entre dos series de resultados es o no significativa a un determinado nivel de confianza.

e) *Ambas curvas normales son simétricas respeto al mismo eje vertical, si bien la curva 1, con desviación estándar menor resulta más puntiaguda (menos aplanada).*



f)

	2				
	2	***3***			
	2	***3***	***4***		
	2	***3***	***4***	***5***	
1	***2***	***3***	***4***	***5***	***6***

moda < mediana < valor medio

Valor medio	*3,1875*
Moda	*2*
Mediana	*3*

g)

					6		*valor medio < mediana < moda*
				5	**6**		
			4	**5**	**6**		*Valor medio 4,4375*
		3	**4**	**5**	**6**		*Moda 6*
1	**2**	**3**	**4**	**5**	**6**		*Mediana 5*

h)

			4				*moda > valor medio > mediana*
		3	**4**				
	2	**3**	**4**				*Valor medio 3,2941*
1	**2**	**3**	**4**	**5**			*Moda 4*
1	**2**	**3**	**4**	**5**	**6**		*Mediana 3*

i)

1							*mediana > valor medio > moda*
1			**4**				
1		**3**	**4**				*Valor medio 2,8125*
1	**2**	**3**	**4**				*Moda 1*
1	**2**	**3**	**4**	**5**	**6**		*Mediana 3*

j) *Fiabilidad es la propiedad por la que un instrumento, un técnico, etc. se comporta de la forma prevista a lo largo del tiempo. Este concepto incluye, simultáneamente, los requisitos de exactitud y precisión exigibles.*

La exactitud es la propiedad por la que una medida o el representante de una serie de ellas se acercan al valor real o de referencia aceptado.

La precisión es una medida del grado de dispersión o concordancia que presentan los resultados obtenidos cuando se mide repetidamente el valor de una variable.

k) *Durante la toma de medidas con un instrumento analítico podemos diferenciar entre:*

-Ruido, que es un señal inevitable de naturaleza aleatoria que presenta el instrumento y se superpone al señal analítico que queremos apreciar. El efecto de este ruido puede disminuir apreciablemente según las condiciones en que se efectúe la medida y, evidentemente, de acuerdo a la calidad de los componentes del instrumental.

-Límite de detección, que corresponde a un determinado valor de concentración del analito, definido matemáticamente por convenio, a partir del que podremos afirmar, con una cierta seguridad estadística, que la señal medida corresponde a la aportación del analito y no únicamente a la del ruido existente.

-Límite de cuantificación, que corresponde a aquel valor de concentración del analito, también definido matemáticamente por convenio, a partir del que es posible la cuantificación del analito.

l) *Sí, por ejemplo, si solo se utilizan dos puntos para la obtención de la recta de calibración. La linealidad es perfecta pero su aplicabilidad muy discutible*

m) *Ambos conceptos son formas diferentes de medir el grado de precisión de una serie de medidas. Si se mantienen las variables durante la adquisición de estas (aparato, técnico, disoluciones, etc.) la medida de la precisión se establece en términos de repetibilidad. Si, contrariamente, cambian las condiciones en la obtención de las medidas (laboratorio,* *instrumentación, técnico...) hablamos entonces de la precisión en términos de reproductibilidad.*

n) *Ambos requisitos previos son necesarios para obtener medidas aceptables. La obtención del cero de un aparato supone un ajuste interno del instrumento, sin emplear ninguna solución o patrón, a partir de reguladores y procedimientos establecidos que modifican convenientemente determinados niveles de señales (mecánicos, ópticos, eléctricos, etc.) propios del aparato. Esta regulación puede ser, según el caso, manual o automática.*

El blanco de un método corresponde a aquella disolución empleada que, en las condiciones de la medida, permite el ajuste de la respuesta del instrumento a un valor determinado (habitualmente el señal analítico cero).

o) *No.*
El ajuste del cero del aparato es independiente del método que se utilice y se consigue regulando convenientemente determinados niveles de señal en ausencia de soluciones de comparación. Hay que proceder a este ajuste del cero del aparato, previamente a la medida de las disoluciones de ensayo, para un método concreto, con las que se podrá fijar, finalmente el cero operativo correspondiente.

Con el cero del aparato, con respecto a un método concreto, medimos un medio, sin el analito, de tal forma que marque cero (o lo establecido como tal) en el procedimiento normalizado de trabajo. Este concepto está relacionado con el concepto de "blanco".

8.3

a) *La molaridad de la disolución se calcula como el número de moles de HCl presentes en 1 litro de disolución. Previamente, calcularemos la masa molecular del ácido.*

$M\ H = 1{,}00794 \pm 0{,}00007$
$M\ Cl = 35{,}453 \pm 0{,}001$
$M\ HCl = (1{,}00794 \pm 0{,}00007) + (35{,}453 \pm 0{,}001) = 36{,}46094 \pm I_A$
$I_A = \sqrt{0{,}00007^2 + 0{,}001^2} = 1{,}0024... \cdot 10^{-3}$
$M\ HCl = 36{,}461 \pm 0{,}001$ g/mol

La incertidumbre correspondiente al matraz aforado de 250 mL es del 0,1%. Por tanto, de forma absoluta:
$0{,}1/100 \cdot 250 = 0{,}25$ mL; $V = 250{,}00 \pm 0{,}25$ mL

$$[HCl] = 1{,}234 \pm 0{,}002\ g \cdot \frac{1\ mol}{36{,}461 \pm 0{,}001\ g} \cdot \frac{1}{250{,}00 \pm 0{,}25\ mL} \cdot \frac{1000\ mL}{1\ L}$$

$[HCl] = 0{,}1353775... \pm I_A$
$I_A = 0{,}1353775... \cdot \sqrt{(0{,}002/1{,}234)^2 + (0{,}001/36{,}461)^2 + (0{,}25/250)^2}$
$I_A = 0{,}1353775... \cdot 1{,}904617... \cdot 10^{-3} = 2{,}578... \cdot 10^{-4} \cong 0{,}0003$

Finalmente,
$[HCl] =$ ***0,1354 ± 0,0003 mol/L***

b) *Suponiendo la disociación completa del ácido, la concentración de protones en la disolución es igual a la concentración inicial del ácido.*
$[H^+] = 0{,}1354 \pm 0{,}0003$ mol/L

El valor de la concentración presenta 4 cifras significativas, por tanto:
$pH = -\log [H^+] = 0{,}8683813...$
$pH =$ ***0,8684***

8.4

a) *Comparamos los valores medios de cada serie con el test t correspondiente. Aceptamos la hipótesis de considerar que las puntuaciones obtenidas corresponden a toda la población (el parámetro de dispersión a considerar será σ y no s)*

Parte	***N***	$\bar{x}$	***s***
Primera parte (fórmula)	*10*	*23,8*	*5,546...*
Segunda parte (nombre)	*10*	*21,2*	*7,859...*

$$s_p^2 = \frac{(N_1 - 1)\, s_1^2 + (N_2 - 1)\, s_2^2}{N_1 + N_2 - 2} = \frac{(9 \cdot 5{,}546...^2) + (9 \cdot 7{,}859...^2)}{18} = 46{,}261...$$

$s_p = 6{,}8015$

$$t_{calc} = \frac{|\bar{x}_1 - \bar{x}_2|}{s_p} \sqrt{\frac{N_1 \cdot N_2}{N_1 + N_2}} = \frac{|23{,}8 - 21{,}2|}{6{,}8015} \sqrt{\frac{100}{20}} = 0{,}855$$

t_{tab} *(95%, 18 gdL) = 2,101*

$t_{calc} < t_{tab}$*, por tanto, no podemos afirmar que para este nivel de confianza la diferencia existente sea significativa.* ***No podemos afirmar que una parte sea más fácil que la otra****.*

Puede comprobarse, como ejercicio complementario, que la conclusión no variaría si hubiésemos considerado los datos como los correspondientes a una muestra.

b) *Calcularemos las correspondientes puntuaciones finales:*

Alumno	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*
Nota	*29*	*30*	*28*	*27*	*26,5*	*24,5*	*16*	*18*	*14*	*11,5*

Para esta serie, $\bar{x} = 22{,}45$; $\sigma = 6{,}517...$; $N = 10$

t_{tab} *(95%, 9 gdL) = 2,2621*

$\Delta = \pm\, \sigma\, t / \sqrt{N} = \pm\, 6{,}517... \cdot 2{,}2621/\sqrt{10} = 4{,}662...$

El valor medio será:

$\bar{x} \pm \Delta = 22{,}45 \pm 4{,}662... =$ ***22 ± 5***

8.5

a) *Aplicaremos el test F*

Método	***N***	$\bar{x}$	***s***
1 (estándar)	*8*	*25,50*	*3,31*
2 (nuevo)	*8*	*28,14*	*2,30*

$$F_{calc} = \frac{s_1^2}{s_2^2} = \frac{3{,}31^2}{2{,}30^2} = 2{,}071...$$

F_{tab} *(95%,* $n_1 = n_2 = 7$ *gdL) = 3,79*

(gdL = grados de Libertad)

$F_{calc} < F_{tab}$*, para este nivel de confianza, la diferencia existente entre las variancias de ambos métodos no es significativa.* ***No podemos afirmar que el método nuevo sea más preciso que el estándar.***

b) *Para $N_1 = N_2 = 50$, cambia el valor de F_{tab}. En este caso, observamos que en la tabla correspondiente no aparece el correspondiente valor. No obstante, sabemos que estará comprendido entre:*

F_{tab} *(95%, $n_1 = n_2 = 40$ gdL) = 1,69*
F_{tab} *(95%, $n_1 = n_2 = 60$ gdL) = 1,53*
F_{tab} *(95%, $n_1 = n_2 = 49$ gdL) $\cong$ 1,6*

En cualquier caso, la conclusión final será la misma ya que F_{tab} resulta inferior a F_{cal}
Así pues, para 50 muestras la diferencia entre las variancias sería significativa y podríamos afirmar que el método nuevo presenta una precisión superior, es decir, ***el método estándar presenta errores aleatorios superiores*** *con respecto al método nuevo.*

8.6

a) *Tabulamos los datos:*

Método	*N*	$\bar{x}$	*s*	s^2
1 (nuevo)	*10*	*25,6*		*3,2*
2 (estándar)	*10*		*1,8*	

Calculamos la variancia correspondiente al método estándar:

$V_2 = s_2^2 = 1,8^2 = 3,24$

$$F_{calc} = \frac{s_2^2}{s_1^2} = \frac{3,24}{3,2} = 1,0125$$ *(Recordemos que F_{calc} no puede ser inferior a 1)*

F_{tab} *(95%, $n_1 = n_2 = 9$ gdL) = 3,18*

Con un nivel de confianza del 95 %, las diferencias existentes entre las variancias de ambos métodos no son significativas. ***No podemos afirmar que ninguno de los dos métodos sea más preciso que el otro.***

b) $\bar{x}$ **=** *25,6; s =* $\sqrt{3,2}$ *= 1,7888*
Tipificamos el valor correspondiente a x = 30 ppm
z = (30 - 25,6) / 1,7888 $\cong$ 2,46

P (x ≥ 30) = P (z ≥ 2,46) = 1 - P (z < 2,46) = 1 - 0,9931 = 0,0069
P (x ≥ 30) = ***0,69%***

c) *Considerando un 40% del área a la izquierda y un 40% del área a la derecha del valor central, el área definida dentro del intervalo [z_1; z_2] dejará en cada uno de los extremos un 10 % del área total:*

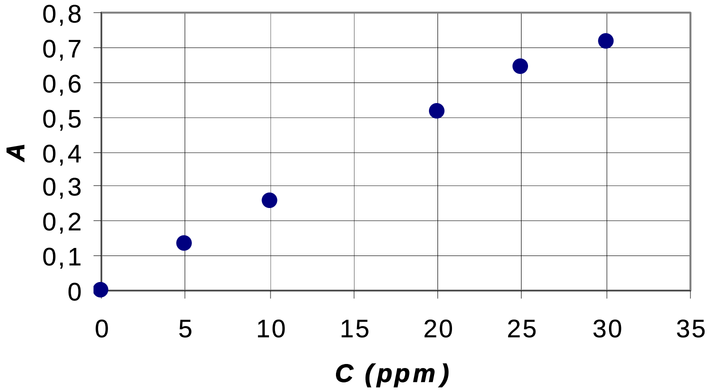
P = 90% = 0,90, en tablas $z_2 \cong 1,28$, ya que P (z ≤ 1,28) $\cong$ 0,90
$z_2 = (x_2 - \bar{x})/s$; $x_2 = \bar{x} + z_2\, s = 25,6 + 1,28 \cdot 1,7888 = 27,9$

Por tratarse de un intervalo simétrico respecto al valor medio, $z_2 = -z_1$
$x_1 = \bar{x} + z_1\, s = \bar{x} - z_2\, s = 25,6 - 1,28 \cdot 1,7888 = 23,3$

El intervalo que contiene el 80 % de los datos centrados sobre el valor medio es:
[23,3 ; 27,9]

8.7

a) *Se muestra la representación correspondiente obtenida con la hoja de cálculo Excel.*



b) *Calcularemos las pendientes correspondientes con la ayuda de la matriz de las pendientes:*

PUNTO	***0***	***5***	***10***	***20***	***25***	***30***
0	---	0,02480	0,02480	0,02475	0,02488	0,02533
5	---	---	0,02480	0,02473	0,02490	0,02544
10	---	---	---	0,02470	0,02493	0,02560
20	---	---	---	---	0,02540	0,02650
25	---	---	---	---	---	0,02760
30	---	---	---	---	---	---

En total hay 15 valores. La posición de la mediana será la octava. Ordenamos las pendientes de menor a mayor y localizaremos el valor que ocupa la posición número 8:

0,02470 - 0,2473 - 0,2475 - 0,02480 -0,02480 - 0,02480 -0,02488 - ***0,02490 -*** *...*
$m^{MS} = 0,02490$

Ahora encontraremos los diferentes valores de la ordenada en el origen:

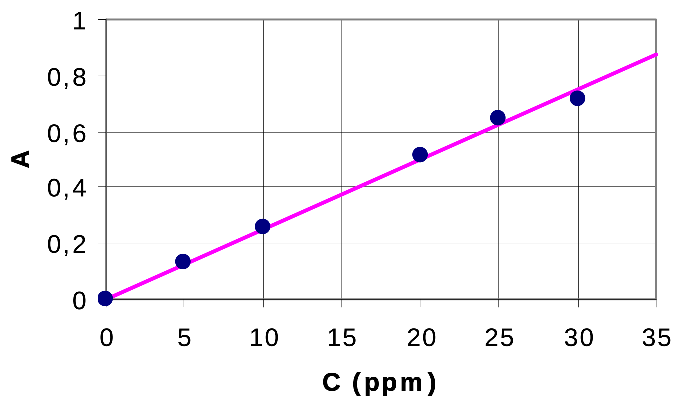
x	***y***	***b = y - 0,02490 x***
0	0,003	0,0030
5	0,127	0,0025
10	0,251	0,0020
20	0,498	0,0000
25	0,625	0,0025
30	0,763	0,0160

Ordenando los valores de b,
0,0000 - 0,0020 - ***0,0025*** *-* ***0,0025*** *- 0,0030 - 0,0100*

La serie anterior de 6 elementos tiene la mediana situada en la posición (6+1)/2 = 3,5, es decir, corresponde al promedio de los valores situados entre la tercera y la cuarta posición. En nuestro caso:
$b^{MS} = 0,0025$

La ecuación resultante será:
y = 0,02490 x + 0,0025

Se muestra la representación de la recta así como los puntos experimentales:



Puede comprobarse como, a pesar de la presencia de valores anómalos, la recta sigue la tendencia general de los datos.

c) *Análisis de los residuales. Los valores correspondientes a y_{calc} se encuentran a partir de la ecuación anterior: $y_{calc} = 0,02490\ x_{exp} + 0,0025$.*

x	y_{exp}	y_{calc}	$\lvert y - y_{cal}\rvert^2$	% SQR
0	0,003	0,0025	$2,50 \cdot 10^{-7}$	0,13
5	0,127	0,1270	0	0,00
10	0,251	0,2515	$2,50 \cdot 10^{-7}$	0,13
20	0,498	0,5005	$6,25 \cdot 10^{-7}$	3,31
25	0,625	0,6250	0	0,00
30	0,763	0,7495	$1,8225 \cdot 10^{-7}$	96,43
		Σ (SQR)	**$1,890 \cdot 10^{-4}$**	**100,00**

Tal como también se observa en la representación gráfica, el punto (30 ; 0,763) es el que más se aleja de la tendencia general. El análisis de residuales muestra que, únicamente este punto, acumula el 96,43 % de la suma del cuadrado de los residuales. Debería descartarse, por tanto, el punto ***(30 ; 0,763)*** *y con el resto proceder al tratamiento por mínimos cuadrados.*

d) *$y = 0,02490\ x + 0,0025$. De acuerdo a las variables estudiadas, la ecuación encontrada en (b) puede ser descrita como:*

$A = 0,02490\ C + 0,0025$
$C = (A - 0,0025) / 0,02490$

Para el valor $A = 0,288$, resulta:
$C = (0,288 - 0,0025) / 0,02490 = 11,4659...$ ppm ≅ ***11,47 ppm***

8.8

a) *La distribución se puede describir como N (0,25 ; 0,08). El 20 % de las preparaciones, para un total de 30, corresponderá a:*
$30 \cdot 20/100 = 6$ preparaciones

El 80 % de las preparaciones no consideradas estarán situadas un 40 % por encima y un 40 % por debajo del intervalo central. El límite superior del intervalo, que dejará a la *izquierda a un 60 % del área corresponde a un valor de z situado entre 0,25 y 0,26 (ver tablas).*

Área	***z***
0,5987	*0,25*
0,6000	z_{sup}
0,6026	*0,26*

Como ejemplo, interpolaremos para definir con más rigor el valor de z_{sup}

$$\frac{z_{sup} - 0,25}{0,6000 - 0,5987} = \frac{0,26 - 0,25}{0,6026 - 0,5987}$$

$z_{sup} = 0,25333... \cong 0,253$

$z_{sup} = (x_{sup} - \bar{x}) / s$

$x_{sup} = \bar{x} + z_{sup} \cdot s = 0,25 + 0,253... \cdot 0,08 = 0,270... \cong 0,27$

Análogamente, por tratarse de un intervalo centrado sobre el valor medio:

$x_{inf} = \bar{x} - z_{sup} \cdot s = 0,25 - 0,253... \cdot 0,08 = 0,22973... \cong 0,23$

El intervalo solicitado de concentraciones será: ***[0,23 ; 0,27]***

b) *Las preparaciones aceptadas están incluidas en el rango [0,20 ; 0,30]. Teniendo en cuenta los parámetros que definen la distribución, N (0,25 ; 0,08), tipificamos los correspondientes valores de x:*

$$z = \frac{x - \bar{x}}{s}$$

$x_{inf} = 0,20 \quad \rightarrow \quad z_{inf} = (0,20-0,25)/0,08 = -0,625$

$x_{sup} = 0,30 \quad \rightarrow \quad z_{inf} = (0,30-0,25)/0,08 = 0,625$

El área solicitada viene determinada por:

$P(x \leq x_{sup}) - P(x \leq x_{inf}) = P(z \leq z_{sup}) - P(z \leq z_{inf})$

Área	***z***
0,7324	*0,62*
---	*0,625*
0,7357	*0,63*

Para z = 0,625 podemos considerar que el área correspondiente es:

(0,7324 + 0,7357) / 2 = 0,7340

$P(z \leq 0,625) = 0,7340$

$P(z \geq 0,625) = 1 - 0,7340 = 0,2660 = P(z \leq -0,625)$

Finalmente,

$P(z_{inf} \leq z \leq z_{sup}) = 2 \cdot 0,2660 = 0,5320 =$ ***53,20%***

8.9

La comparación de precisiones se efectúa a partir del test F. Con la ayuda de la calculadora científica, encontramos los correspondientes valores de desviación estándar:

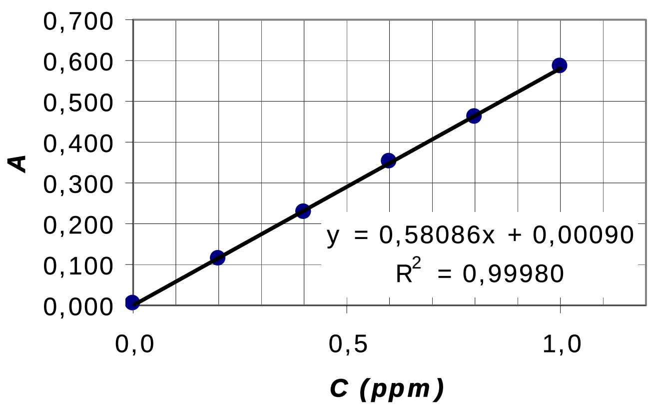
Método	***N***	***s***
Absorción	*7*	*7,8831...*
Fluorescencia	*7*	*6,5574...*

$F_{calc} = (7{,}8831...)^2/(6{,}5574...)^2 = 1{,}445...$
F_{tab} *(95 %,* $n_1 = n_2 = 6$ *gdL) = 4,28*

$F_{calc} < F_{tab}$*, en las condiciones de la comparación no hay diferencia significativa entre las variancias de las dos series.* ***No podemos afirmar que una de las técnicas sea menos precisa que la otra y podemos considerar que presentan precisiones similares.***

8.10

a, b) *Se muestra la representación de la calibración obtenida con una hoja de cálculo, así como la ecuación de regresión por mínimos cuadrados:*



c) *La propia hoja de cálculo permite encontrar el coeficiente de correlación:*
$r^2 = 0{,}99980 \rightarrow$ ***r = 0,9999***

Los 4 nueves del coeficiente permiten calificar el grado de ajuste como "muy bueno"

d) *Por sustitución en la ecuación encuentro:*
$A = 0{,}58086\ C + 0{,}00090$
$A = 0{,}378$

$C = (0{,}378 - 0{,}00090) / 0{,}58086 = 0{,}6492... \cong$ ***0,649 ppm***

8.11

a) *Comparamos las variancias de cada serie*

$F_{calc} = 25{,}40 / 20{,}21 = 1{,}2568...$

Considerando que el número de grados de libertad para la serie del numerador y denominador son, respectivamente, n_1 *y* n_2*, el valor tabulado a emplear será:*

F_{tab} *(95 %,* $n_1 = 39$*,* $n_2 = 49$*)*

En la tabla, no aparece este valor:

	n_1	
n_2	**30**	**40**
40	1,74	1,69
60	1,70	1,59

Es fácil deducir que debe estar alrededor de 1,60 y que, en cualquier caso, será superior al valor de F_{calc}

$F_{calc} < F_{tab}$, al nivel de confianza seleccionado no se puede establecer que la diferencia entre las variancias correspondientes a las dos series sea significativa. ***No podemos afirmar que la variabilidad del grupo femenino sea superior a la del grupo masculino.***

b) *Comparemos los valores medios de cada serie.*

$$s_p^2 = \frac{(N_1 - 1)\, s_1^2 + (N_2 - 1)\, s_2^2}{N_1 + N_2 - 2} = \frac{49 \cdot 20{,}21 + 39 \cdot 25{,}40}{98} = 20{,}213\ldots$$

$s_p = 4{,}496\ldots$

$$t_{calc} = \frac{|\bar{x}_1 - \bar{x}_2|}{s_p} \sqrt{\frac{N_1 \cdot N_2}{N_1 + N_2}} = \frac{|32{,}1 - 33{,}2|}{4{,}496\ldots} \sqrt{\frac{2000}{90}} = 1{,}153\ldots$$

t_{tab} *(95%, 88 gdL)* $\cong 1{,}99$
(El valor en concreto no consta en la tabla y se trata de una apreciación)

$t_{calc} < t_{tab}$, por tanto, no podemos afirmar que a este nivel de confianza la diferencia existente entre los valores medios sea significativa. ***No podemos afirmar, a partir de los resultados de este test, que el sexo tenga influencia en el nivel de atención - percepción.***

8.12

Las diferencias relacionadas con las frecuencias se evalúan con la prueba de xi cuadrado. Consideramos, como hipótesis, que la frecuencia de aparición de los dígitos es equiprobable:

Dígito	**FO**	**FE**	**FO - FE**	**$(FO\text{-}FE)^2/FE$**
0	200	205,6	-5,6	0,15253
1	212	205,6	6,4	0,19922
2	195	205,6	-10,6	0,54650
3	220	205,6	14,4	1,00856
4	235	205,6	29,4	4,20409
5	185	205,6	-20,6	2,06401
6	190	205,6	-15,6	1,18366
7	205	205,6	-0,6	0,00175
8	210	205,6	4,4	0,09416
9	204	205,6	-1,6	0,01245
Σ	**2056**	**2056**	**0,0**	**9,46693**

(FO = Frecuencia observada, FE = Frecuencia esperada)
$\chi^2_{calc} = 9{,}46693\ldots$
χ^2_{tab} *(95%, n = 9 gdL) = 16,92*

Como $\chi^2_{calc} < \chi^2_{tab}$, la diferencia existente entre les frecuencias observadas y las esperadas no es significativa. ***No podemos afirmar que exista ninguna causa que haga aparecer unos dígitos más que otros en la última posición.***