

PREGUNTAS SELECCIONADAS DE EXÁMENES ANTERIORES

Diferencie entre errores de tipo I (α) y errores de tipo II (β). Indique dos situaciones características en que sea más peligroso cometer un tipo u otro.

Uno de nuestros proveedores afirma que el producto que nos vende (carbonato cálcico a utilizar como blanqueante) tiene una pureza del 98,50 % y una desviación típica $\sigma = 1,27$ %. Para comprobarlo, nuestro Departamento de Control de Calidad analiza 15 porciones de material, encontrando una pureza del 96,43 %. Indique como se puede enfocar quimiométricamente el problema y determine si deberíamos llamar la atención del proveedor.

¿Cuáles son los parámetros y términos que aparecen en una tabla de ANOVA de una vía y cómo se obtienen?

Indique claramente mediante dos ejemplos cuando se puede hablar de regresión y cuando de correlación entre dos variables. ¿Qué es la falta de ajuste de un modelo lineal y como puede detectarse?

¿Qué es un bias en un método analítico y cuantas clases hay de ellos?. ¿Qué es la precisión, cómo se cuantifica y cuántas clases hay?. Esquematice gráficamente cuáles de los anteriores componentes pueden determinarse en un estudio en el que participen varios laboratorios

Teorema del límite central. Enunciados. Intervalos de confianza de la media. Muestras pequeñas y distribución t

¿Qué es el método rankit y para qué se utiliza?. Esquematice sus etapas.

Interés y aplicaciones de las pruebas de normalidad en Quimiometría. Enumere las etapas de la prueba χ^2 .

Se desea conocer si la agitación influye la determinación del pH de muestras de suelos, para lo cual se mide el pH de 20 suelos diferentes de suelo con y sin agitación, originando una tabla de 20 filas por 2 columnas. Indique claramente cómo podría avaluar la existencia de dicho efecto y qué etapas deben seguirse para ello.

Justifique gráficamente qué bias puede estimar un laboratorio que trabaja en solitario. ¿Cómo se pueden estimar los bias si el laboratorio analiza muestras con intervalo de concentraciones amplio y material blanco disponible?. ¿Y cuando no se dispone de material blanco?

Se han determinado los contenidos en potasio extraíble en los dos horizontes edafológicos (A y B) de una serie de 5 muestras de 5 suelos agrícolas.

Muestra	K (mg / g)	
	A	B
S132	0,55	0,54
S144	0,61	0,52
S171	0,99	0,83
S205	0,62	0,58
S206	1,12	0,81

Demuestre si existen diferencias significativas entre los dos horizontes.

Indique y comente brevemente al menos dos (y mejor si son tres) procedimientos que permitan detectar si una muestra tiene una distribución normal o no.

Sospechamos que uno de nuestros proveedores nos proporciona un producto cuyos niveles en cadmio son superiores a nuestras especificaciones (10,00 $\mu\text{g/kg}$). Para ello nuestro Departamento de Calidad analiza 20 porciones de un cargamento del producto suministrado, encontrando un valor medio de cadmio de 10,20 mg/kg con una desviación típica de 0,50

mg/kg. Indique como se puede enfocar el problema y determine si nuestras sospechas son ciertas con un nivel de significación del 5%.

¿Para qué sirve el ANOVA de una vía?. Describa brevemente los componentes de la tabla del ANOVA explicando cómo se reparten los términos que aparecen en ella. ¿Cuál es la prueba final y cómo se explican sus resultados?. ¿Qué significa en ANOVA de una vía que el p "a posteriori" sea mayor de 0,05?

¿Para qué sirve el ANOVA de dos vías?. Indique cómo se reparten los términos correspondientes a las sumas de cuadrados. Resuma y describa la tabla del ANOVA correspondiente.

Un laboratorio analítico dispone de 5 analistas de laboratorio que utilizan un mismo procedimiento estándar para analizar cadmio en muestras de tejidos biológicos. Se ha observado que al analizar por sextuplicado un mismo material, los 5 analistas obtienen los siguientes resultados:

Analista				
1	2	3	4	5
5,22	5,28	5,58	5,35	5,35
5,63	5,56	5,50	5,35	5,35
5,35	5,30	5,50	5,35	5,35
5,19	5,28	5,52	5,35	5,35
5,12	5,20	5,46	5,32	5,32
5,29	5,31	5,50	5,34	5,34

Un Análisis de Varianza origina la siguiente tabla que debe ser rellenada

Fuente	SS	g.d.l.	MS	F _{calculado}	F _{crítico}
Entre grupos (A)	0,1711				2,759
Dentro grupos (R)					
Total	0,4138				

¿Qué significa cada uno de los términos que aparecen en el encabezamiento de la tabla del ANOVA?. ¿Qué conclusiones pueden deducirse?. ¿Es el modelo de efectos fijos o aleatorios?. ¿Qué significa el valor de p "a posteriori" si resulta que vale $p=0,0079$?

¿Cuál es la diferencia entre regresión y correlación?. ¿Qué significa el llamado comúnmente coeficiente de regresión r que proporcionan las calculadoras?. ¿Qué son elementos espurios en un Análisis por Regresión Lineal?. ¿Cómo pueden detectarse?

Se está adquiriendo un determinado reactivo para análisis que debería contener como máximo (de acuerdo con el fabricante) $520 \mu\text{g.kg}^{-1}$ ($\sigma = 35 \mu\text{g.kg}^{-1}$). Se analizan 10 porciones de reactivo, encontrando $\bar{x} = 563 \mu\text{g.kg}^{-1}$. ¿Debe rechazarse el material?. Si se descubre que nuestro método analítico tiene un bias de manera que la media de nuestras determinaciones del material es de $580 \mu\text{g.kg}^{-1}$, determine gráficamente la manera de encontrar probabilidad de cometer un error de tipo II.

Etapas de un test de hipótesis. ¿Qué es un test de una cola y de dos colas y cuando deben ser empleados?.

¿Qué es un modelo de efectos fijos y qué un modelo de efectos aleatorios en ANOVA?. Indique algún ejemplo.

¿Qué son los outliers en ajuste lineal?. Diferencie, mediante algún gráfico entre los diversos tipos de outliers.

¿Qué significa en regresión lineal que el intervalo de confianza de la ordenada en el origen incluya a cero?.

¿Tiene algún significado el llamado "coeficiente de regresión" r que se obtiene al hacer un ajuste por mínimos cuadrados con una calculadora?.

Durante un proceso de validación de un suelo (que se pretende comercializar como Material de Referencia Certificado para la determinación de P extraíble en HCl) se analizan 4 muestras por triplicado, en condiciones homogéneas, obteniendo los siguientes resultados en mg/kg

BO1291	BO1292	BO1293	BO1294
39,3	38,7	38,9	39,3
40,6	39,9	39,8	38,7
39,9	39,7	39,7	39,1

El Análisis de Varianza origina la siguiente tabla que debe ser rellenada

Fuente	SS	g.d.l.	MS	F _{calculado}	F _{crítico} ($\alpha=0,05$)
Entre grupos (A)					4,07
Dentro grupos (R)	2,347				
Total	3,567				

- ¿Qué quiere decir condiciones homogéneas?
- ¿Qué significa cada uno de los términos que aparecen en el encabezamiento de la tabla del ANOVA?
- Se obtiene además la siguiente información:

Modelo de efectos fijos	
t	2,3063
LSD	1,020

Modelo de efectos aleatorios	
s_A^2	0,0378
s_R^2	0,2933

¿Es un modelo de efectos fijos o aleatorios?. ¿Cuál es el parámetro característico a estudiar en este caso?

- ¿Qué significa que el valor de p “a posteriori” sea 0,315?

Defina y diferencie entre límite de detección, límite de decisión y límite de cuantificación de un procedimiento analítico. ¿Cómo se pueden estimar los bias de un procedimiento analítico que se está validando cuando el rango de concentraciones de analito es amplio y no se dispone de un material en blanco?

¿Qué tres suposiciones implícitas hace el método de ajuste lineal por mínimos cuadrados?. ¿Qué son los residuales?. ¿Cómo se pueden utilizar para determinar la bondad del ajuste y cuáles son las situaciones deseables e indeseables más frecuentes?

¿Cuál de las dos líneas de calibrado siguientes sería la más adecuada para el estudio de la falta de ajuste del modelo lineal en un nuevo método analítico y por qué?

- Línea 1 preparada con disoluciones que contienen 0,0; 1,0; 2,0; 3,0; 4,0; 5,0; 6,0; 7,0; 8,0; 9,0; 10,0
- Línea 2 preparada con disoluciones que contienen: 0,0; 2,0; 2,0; 4,0; 4,0; 6,0; 6,0; 8,0; 8,0; 10,0; 10,0 (concentraciones en $\mu\text{g.l}^{-1}$)

Diferencie entre errores α y errores β en una prueba estadística de significación. Indique cómo están relacionados entre sí. ¿Qué es la potencia de una prueba o test estadístico?

¿Cuáles son los parámetros y términos que aparecen en una tabla de ANOVA de una vía y cómo se obtienen?. Indique un esquema de trabajo que permita evaluar simultáneamente el efecto que tienen sobre una determinación analítica el cambio de analista, instrumento y día de la semana.

¿Qué es el ANOVA encajado (nested ANOVA) y para qué sirve?. Aplíquelo al caso de un laboratorio que dispone de 3 analistas y dos instrumentos diferentes para la determinación de la conductividad de aguas de bebida.

Se desea evaluar una recta de calibrado construida mediante el siguiente diseño experimental:

C_i (mg/l)	0	1	2	3	4	5
Réplicas	2	2	2	2	2	2

obteniéndose esta tabla (incompleta) de ANOVA

Fuente	g.d.l	SS	MS	F
Regresión		20,310		
Residual				
Falta de ajuste		0,164		
Puro error				
Total		20,792		

- 1) Complete las casillas sombreadas de la tabla del ANOVA
- 2) Indique si el modelo explica una cantidad significativa de información
- 3) Determine si hay faltas de ajuste
- 4) Calcule el R^2 .

Proponga dos situaciones reales diferentes, una en la que resulte más peligroso cometer un error α (o de tipo I) y otra en la que sea más peligroso un error β (o de tipo II). Demuestre cómo se puede calcular el tamaño n de la muestra (n° de réplicas) que debe ser analizado para poder detectar una diferencia (bias) δ , mediante un procedimiento que tiene una desviación típica σ , de manera que estemos protegidos simultáneamente frente a los errores de tipo α y β .

Se desea evaluar una recta de calibrado construida mediante el siguiente diseño experimental:

C_i (mg/l)	0,0	0,1	0,2	0,3	0,4	0,5
Réplicas	2	3	2	2	3	2

por medio de una ANOVA, cuya tabla (incompleta) es:

Fuente	g.d.l	SS	MS	F
Regresión		10,792		
Residual				
Falta de ajuste		0,218		
Puro error				
Total		11,174		

- 5) Complete las casillas sombreadas de la tabla del ANOVA
- 6) Indique si el modelo explica una cantidad significativa de información
- 7) ¿Hay falta de ajuste en el modelo?
- 8) ¿Cuánto vale R^2 y qué significa?

¿Qué son los residuales en un análisis de regresión?. Dibuje esquemáticamente las formas más habituales que pueden tener los gráficos de residuales. ¿Qué es la falta de ajuste de un modelo lineal y cómo puede detectarse de forma no gráfica?. Indique claramente mediante dos ejemplos qué es regresión y qué es correlación entre dos variables.

Se dispone de dos procedimientos diferentes para la determinación de carbonatos en suelos (calcímetro y valoración ácido-base) y se desea conocer si ambos dan resultados equivalentes. Para ello se analizan 7 suelos diferentes por ambos métodos obteniendo los resultados que se reflejan en la tabla siguiente.

Método	Muestra						
	1	2	3	4	5	6	7
Calcímetro	17,45	12,32	18,76	11,22	15,88	14,13	16,23
Valoración	17,77	12,52	18,95	11,36	16,24	14,21	15,96

Demuestre si ambos métodos originan o no resultados diferentes con un nivel de significación del 5%. ¿Cómo sería el valor de p a posteriori? (opcionalmente si sabe como calcularlo hágalo)

¿Para qué sirve el ANOVA de una vía?. Describa brevemente los componentes de la tabla del ANOVA explicando cómo se reparten los términos que aparecen en ella. ¿Cuál es la prueba final y cómo se explican sus resultados?. ¿Qué significa en ANOVA de una vía que el p a posteriori sea mayor de 0,05?

Defina (con la ayuda de esquemas gráficos) los conceptos de Límite de decisión (LC) y límite de detección (LD) suponiendo que todas las distribuciones implicadas son normales y tienen una varianza σ^2 idéntica.

Nuestra Empresa química ha licenciado una franquicia para que otra Empresa **B** fabrique nuestro producto estrella, y se desea evaluar si dicha Empresa **B** lo fabrica con la pureza requerida. Para ello, un laboratorio acreditado e independiente analiza mediante el mismo procedimiento analítico, dos lotes de producto procedentes de las dos Empresas obteniendo los siguientes resultados en mg/kg:

	Media	s	n
Nuestra Empresa	17,52	0,102	5
Empresa B	17,11	0,104	5

Demuestre, con un nivel de significación $\alpha=0,05$, si la Empresa **B** se adecúa a los requerimientos del contrato, es decir fabrica un producto indistinguible del nuestro.

¿Para qué sirve el ANOVA encajado (nested ANOVA)?. Aplíquelo al caso de un laboratorio que dispone de tres analistas y dos calcímetros diferentes para la determinación de carbonato en suelos.

Describa al menos dos aplicaciones de la Regresión lineal en la Validación de Métodos Analíticos indicando claramente qué valores deben tener los parámetros de las correspondiente rectas de regresión en cada aplicación.

¿Cómo se puede detectar mediante Regresión Lineal la existencia de efecto de matriz en un procedimiento analítico?. Describa al menos dos aplicaciones de la Regresión lineal en la Validación de Métodos Analíticos indicando claramente para cada una de ellas: (a) cuáles son los modelos subyacentes (b) qué valores teóricos deben tener los parámetros de los mismos y (c) qué significa si los parámetros obtenidos se apartan de los teóricos.

Nos han propuesto la posibilidad de sustituir un procedimiento analítico, **A**, que tenemos puesto a punto y validado por otro nuevo, **B**, que resulta más barato. Dentro del proceso de decisión se nos plantea conocer si el nuevo procedimiento **B** origina resultados comparables al antiguo **A**. Para ello se analizan 8 muestras diferentes por ambos procedimientos **A** y **B** obteniendo los resultados que se reflejan en la tabla siguiente.

Procedimiento	Muestra							
	1	2	3	4	5	6	7	8
A	5.80	4.12	6.56	3.15	4.61	5.62	6.13	5.23
B	5.44	4.26	5.51	3.43	4.30	5.77	5.90	4.56

Demuestre si ambos métodos originan o no resultados diferentes con un nivel de significación del 5%. ¿Qué significado tiene el valor de p a posteriori y cuanto valdría en este caso?

Dentro del desarrollo de un nuevo método analítico, se desean comparar dos procedimientos de digestión previos a la determinación del analito. Para ello se analizan diferentes porciones de un material homogéneo aplicando los dos procedimientos y obteniendo los siguientes resultados:

	Procedimiento	
	1	2
$\bar{x}^*$	39.65	42.36
s^*	0.94	0.57
n	9	8
$^* \text{ mg.kg}^{-1}$		

Demuestre con un nivel de significación $\alpha = 0,05$ y utilizando las tablas adjuntas, si ambos procedimientos de digestión originan o no resultados comparables.

¿Para qué sirve el ANOVA encajado (nested ANOVA)? Aplíquelo al caso de un laboratorio que dispone de dos analistas y tres conductivímetros diferentes para la determinación de la conductividad de aguas subterráneas.

Defina (con la ayuda de esquemas gráficos) los conceptos de Límite de decisión (LC) y límite de detección (LD) suponiendo que todas las distribuciones implicadas son normales y tienen una varianza σ^2 idéntica. ¿Qué es el Límite de cuantificación (LQ)? Indique cómo se pueden estimar los límites a partir de los parámetros del ajuste de una línea de calibrado.

Describe dos aplicaciones de la Regresión lineal en la Validación de Métodos Analíticos indicando claramente para cada una de ellas: (a) cuáles son los modelos subyacentes (b) qué valores teóricos deben tener los parámetros de los mismos y (c) qué significa si los parámetros obtenidos se apartan de los teóricos.

Uno de los métodos para validar un modelo de regresión es verificar que el modelo elegido describe adecuadamente la relación entre las dos variables (no hay falta de ajuste) que se hace por medio de un ANOVA. ¿Cuál es el otro método y como puede llevarse a cabo? Se dispone de la siguiente tabla incompleta de un ANOVA de los indicados anteriormente aplicado a una línea de calibrado con la siguiente experimentación.

C_i de analito (µg / kg)	0,0	0,5	1,0	1,5	2,0	2,5
Nº Réplicas	2	1	3	3	1	2

Indique el significado de cada término de la tabla del ANOVA, rellene las casillas incompletas y deduzca qué conclusiones se pueden extraer

Fuente	g.d.l	SS	MS	F
Regresión		20,310		
Residual				
Falta de ajuste		0,264		
Puro error				
Total		20,792		

¿Cuál es el valor del R^2 y qué significa?

¿Qué es un bias y cuántos tipos hay en un método analítico?. ¿Qué es la precisión, cómo se cuantifica y cuántas clases hay?. Esquematice gráficamente cuáles de los anteriores componentes pueden determinarse en un estudio en el que participen varios laboratorios analizando un mismo material mediante un mismo procedimiento

Estamos desarrollando un nuevo método analítico y deseamos conocer el efecto que sobre los resultados finales tienen **dos** factores diferentes: Temperatura (**4** posibilidades o niveles) y Agitación (**3** posibilidades o niveles). Se llevan a cabo **4** réplicas de cada situación y se obtiene la siguiente tabla (incompleta) del ANOVA:

Resultados del ANOVA

Fuente	Sumas de cuadrados	g.d.l.	Cuadrados medios	F	F_{crit}	P a posteriori
Temperatura	78,13					
Agitación	6,46					
Interacción	6,96					
Residual						
Total	128,19					

- 1) ¿Cuántos experimentos hemos llevado a cabo?
- 2) Describa brevemente el significado de los componentes de la tabla de resultados del ANOVA (en **negrilla**) explicando claramente cómo se reparten los términos que pueden repartirse y cómo se calculan los que pueden calcularse.
- 3) ¿Qué es la interacción en este tipo de ANOVA y qué significado tendría en este caso concreto?
- 4) Rellene todas las casillas vacías que pueda (**NO lo haga sobre este enunciado**).
- 5) ¿Cómo debería ser el p a posteriori en cada caso?
- 6) ¿Qué conclusiones se pueden sacar acerca de los efectos estudiados?. (Utilice la Tabla del reverso)
- 7) ¿Cómo se actúa en el caso de que la interacción no sea significativa?. Si fuera el caso, rehaga la tabla del ANOVA y reformule las conclusiones que sacó en el punto anterior.

¿Cómo se puede detectar mediante Regresión Lineal la existencia de efecto de matriz en un procedimiento analítico?. Describa al menos dos aplicaciones de la Regresión lineal en la Validación de Métodos Analíticos indicando claramente para cada una de ellas: (a) cuáles son los modelos subyacentes (b) qué valores teóricos deben tener los parámetros de los mismos y (c) qué significa si los parámetros obtenidos se apartan de los teóricos.

Durante la validación de un nuevo procedimiento para la determinación de plomo(II) en suelos se ha analizado un mismo Material de Referencia Certificado mediante nuestro procedimiento (**X**) y mediante uno ya validado y exento de bias (**Referencia**) obteniendo los siguientes resultados (en mg/kg) :

Procedimiento	
X	Referencia
46,4	44,7
45,7	43,7
47,3	45,2
50,2	46,3
47,5	46,7
50,6	42,8
49,2	
50,6	
45,9	
48,3	

Demuestre si el nuevo Procedimiento **X** origina resultados comparables a los del Procedimiento de **Referencia** ($\alpha = 0,05$). ¿Qué es el p a posteriori y cómo sería su valor en este caso comparado con el nivel de significación?

En el caso de que **X** no diera resultados comparables con los originados por **Referencia** ¿serían aquéllos significativamente mayores o menores que los de este último?

Se diseña un estudio inter-laboratorio en el que un material homogéneo se reparte entre **a** laboratorios en cada uno de los cuales se hacen **b** determinaciones analíticas con **n** réplicas. Diseñe un ANOVA encajado (*nested ANOVA*) para dicho ejercicio. ¿Qué componentes pueden estimarse a partir de los diferentes MS?

1.- Se desea validar un procedimiento analítico, **A**, para la determinación de cinc en aguas potables frente a un procedimiento, **B**, ya contrastado y normalizado. Para ello se aplican ambos procedimientos a 9 muestras diferentes de efluentes industriales, obteniéndose los siguientes resultados en $\mu\text{g.l}^{-1}$.

Procedimiento	
A	B
212,0	185,0
251,9	240,7
286,4	289,8
387,0	370,0
407,0	405,3
255,3	222,2
797,0	667,0
702,0	604,0
312,0	285,0

¿Originan ambos procedimientos resultados comparables con un nivel de significación $\alpha = 0,05$? ¿Qué es el p a posteriori y cómo sería su valor en este caso comparado con el nivel de significación? En el caso de que los resultados no fueran comparables, ¿son los de uno significativamente mayores que los del otro?

1.- Se sospecha que el laboratorio de uno de nuestros proveedores de servicios está dando resultados por defecto, es decir inferiores a los valores correctos. Para comprobarlo, se envían a dicho laboratorio 9 muestras diferentes (numeradas de 1 a 9) que son enviadas también a un laboratorio acreditado de alta calidad (y precio) y cuyos resultados están exentos de sesgo o bias. A la vista de los resultados que aparecen en la siguiente tabla, indique si es cierta, con un nivel de significación $\alpha=0.05$, nuestra sospecha de que el proveedor está dando resultados por defecto

Muestra	Laboratorio sospechoso	Laboratorio acreditado
1	12,8	13,1
2	11,5	11,8
3	13,8	13,9
4	14,2	14,5
5	10,6	10,3
6	9,7	9,9
7	11,5	11,9
8	14,1	14,2
9	9,6	9,7

¿Cómo será el valor de p a posteriori?

1.- Un laboratorio dispone de dos procedimientos analíticos diferentes (**A** y **B**) para llevar a cabo la determinación del contenido en carbonatos en suelos. Se desea comprobar si ambos procedimientos originan resultados comparables aplicándolos a una misma muestra. Durante el estudio se obtienen los siguientes resultados (expresados como porcentaje de CaCO_3)

Procedimiento	
A	B
21,69	20,84
20,72	21,29
22,24	22,48
23,27	21,37
23,19	20,62
23,73	20,84
21,76	21,11
23,09	19,65
20,91	20,5

¿Originan ambos procedimientos resultados comparables con un nivel de significación $\alpha = 0,05$? ¿Qué es el p a posteriori y cómo sería su valor en este caso comparado con el nivel de significación?. En el caso de que los resultados no fueran comparables, ¿son los de uno significativamente mayores que los del otro?