

# Aplicación de Microsoft Excel a la Química Analítica: validación de métodos analíticos

José Marcos Jurado

Departamento de Química Analítica

1 de abril de 2008

## 1. Estadística básica.

### 1.1. Concepto de población y muestra

En estadística, se define población como el conjunto de individuos portadores de información del fenómeno que se estudia. Una muestra sería un conjunto de individuos representativos de dicha población, conjunto que permitiría estudiar el fenómeno sin pérdida significativa de información. Así, si se pretende analizar el contenido en sulfatos de un lago, la población sería el propio lago y la muestra sería una porción de agua representativa del mismo. De este modo, a partir del análisis de la muestra, debe ser posible conocer el contenido en sulfatos de la población.

Cuando nos referimos a datos obtenidos experimentalmente en el laboratorio, la población se refiere a la totalidad de medidas posibles, mientras que la muestra sería un conjunto de estas medidas.

### 1.2. Media aritmética.

Se define la media aritmética como la suma de una serie de individuos dividida por el número de individuos considerados. En el caso de una población, es decir, cuando realizásemos una serie infinita de medidas obtendríamos la media de la población, generalmente designada por  $\mu$ . Cuando se realice una serie de medidas representativas de la población se debe hablar de media muestral. La media muestral se calcula como:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (1)$$

siendo  $x_i$  cada uno de los valores obtenidos y  $n$  el número de valores considerados.

### 1.3. Desviación estándar.

La desviación estándar es un parámetro que da cuenta de la dispersión de los datos obtenidos. En el caso de una población hablaremos de la desviación estándar de la población,  $\sigma$ . Para una muestra se calcula con la siguiente fórmula:

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

Al cuadrado de la desviación estándar se le denomina varianza ( $s^2$ ). A partir de la desviación estándar se puede también calcular el coeficiente de variación o desviación estándar relativa, que es una medida del error relativo y se usa para comparar resultados medidos en distintas escalas.

$$CV = RSD = \frac{s}{\bar{x}} \cdot 100 \quad (3)$$

### 1.4. Distribución de medidas repetidas.

Cuando realizamos medidas de una población, en ausencia de errores sistemáticos, éstas pueden tomar cualquier valor en torno al valor verdadero o media de la población, dependiendo de la desviación estándar de la misma (del error aleatorio puro). Si se dispusiera de todas las medidas y se representasen las frecuencias con que aparece cada valor, frente a los propios valores, se podría ver la distribución de los datos en torno a la media de la población. En las condiciones mencionadas se obtiene una distribución normal o gaussiana como la de la figura.

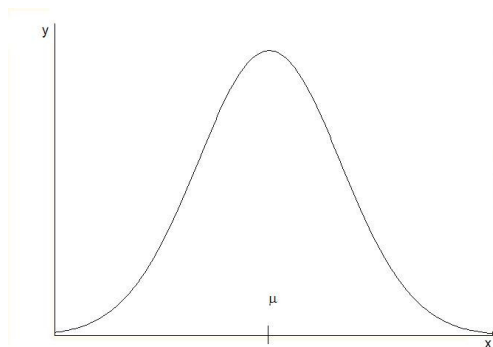


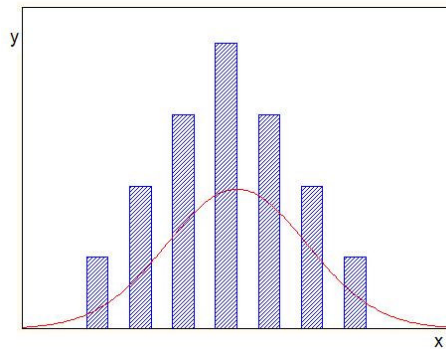
Figura 1. Distribución normal

La distribución normal viene descrita por la ecuación

$$y = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \quad (4)$$

En una distribución normal de media  $\mu$  y desviación estándar  $\sigma$ , aproximadamente el 68 % de los valores de la población caen dentro del intervalo  $\mu \pm 1\sigma$ , el 95 % dentro de  $\mu \pm 2\sigma$  y el 99.7 % dentro de  $\mu \pm 3\sigma$ .

En realidad solo vamos a realizar una serie finita de medidas (replicados) a partir de las cuales se obtiene la media de la muestra y su desviación estándar. La media de la muestra será una estimación de  $\mu$  y la desviación estándar de la muestra será una estimación de  $\sigma$ . Por lo tanto, la distribución de esta muestra será también normal y se cumplirán las mismas características anteriormente mencionadas. Gráficamente, donde antes obteníamos una curva ahora obtendremos un histograma.



Si dispusiésemos de una serie de medias, cada una obtenida de una serie de medidas individuales, se observaría que estas medias se encuentran más agrupadas en torno a  $\mu$  que las medidas individuales. La distribución de todas las medias muestrales posibles se denomina distribución muestral de la media o distribución en el muestreo de la media. Su media es la misma que la media de la población original. Su desviación estándar se denomina error estándar de la media (e.e.m.). Existe una relación matemática exacta entre el e.e.m. y la desviación estándar,  $\sigma$ , de la población:  $e.e.m. = \sigma/\sqrt{n}$ . A mayor  $n$ , menor será el e.e.m. y menor la dispersión de las medias muestrales en torno a  $\mu$ . El e.e.m. nos da una medida de la variabilidad de  $\bar{x}$ , además, aunque la distribución de las medidas no sea una distribución normal, la distribución de las medias muestrales tiende a serlo a medida que aumenta  $n$ . Esto se conoce como la teoría del Límite Central. Se asume que las medias de muestras muy pequeñas ( $<5$ ) se distribuyen normalmente.

### 1.5. Límites de confianza de una medida.

Es poco probable que obtengamos una media de la muestra exactamente igual a la de la población. Generalmente se tiene un intervalo centrado en la media de la muestra que contendrá al valor verdadero con un determinado nivel de confianza. La amplitud de este intervalo depende de dos factores:

- a) La precisión de los valores individuales, que dependen de la desviación estándar de la población.
- b) El número de medidas de la muestra, pues el hecho de repetir las medidas implica tener más confianza en el resultado.

El límite de confianza de una media viene dado por la expresión:

$$\bar{x} \pm z_s / \sqrt{n} \quad (5)$$

Donde, en ausencia de errores sistemáticos:

$$\bar{x} - z_s / \sqrt{n} < \mu < \bar{x} + z_s / \sqrt{n} \quad (6)$$

Donde  $z=1.96$  para un 95% de nivel de confianza, 2.58 para el 99% y 2.97 para el 99.7%.

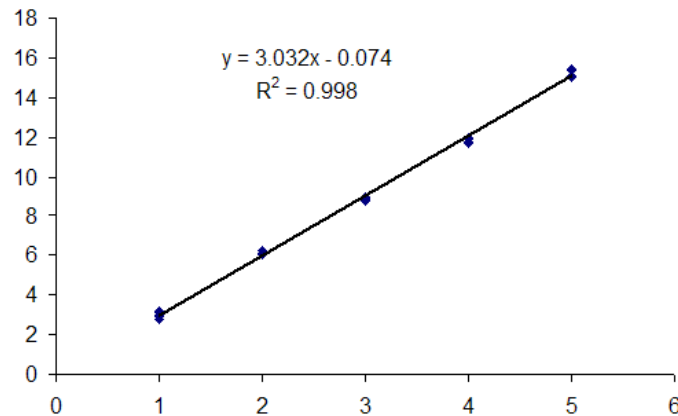
### 1.6. Mediana

Para un conjunto de  $n$  datos ordenados de menor a mayor, la mediana es el valor de la observación que ocupa el lugar  $(n+1)/2$  si  $n$  es impar o bien la media de los valores que ocupan los lugares  $n/2$  y  $(n/2+1)$  si  $n$  es par.

### 1.7. Calibrado.

Supongamos que pretendemos realizar la determinación de un determinado analito mediante un método instrumental, de modo que medimos una propiedad física de la muestra que sea directamente proporcional a la concentración de analito. En principio, a partir de la medida de esta propiedad no podremos calcular la concentración de analito. Para poder hacer esto es preciso establecer una relación entre las señales medidas y una serie de patrones de concentración conocida. Esto es lo que se denomina realizar un calibrado. La relación señal-concentración vendrá dada por una función matemática a partir de la cual se puede interpolar el valor de señal obtenido para una muestra y calcular así la concentración de analito en la misma. Generalmente se procura que la relación sea lineal, lo que simplifica mucho los cálculos a realizar.

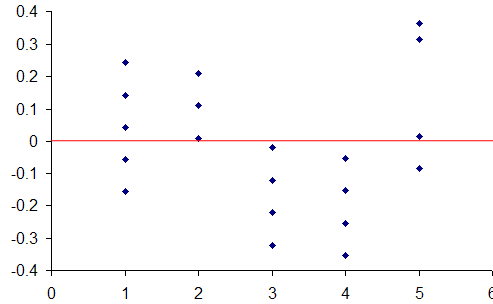
Supongamos que disponemos de una serie de patrones de concentración  $x_1, x_2, \dots, x_i, \dots, x_n$  que, una vez medidos por una determinada técnica presentan unas señales analíticas  $y_1, y_2, \dots, y_i, \dots, y_n$ . Si representamos las señales frente a las concentraciones obtenemos la denominada curva de calibración. En el caso más sencillo, como ya se ha dicho, tendremos una recta de calibrado de ecuación  $y = bx + a$ , donde  $b$  es la pendiente y  $a$  la ordenada en el origen.



Para calcular el valor de estos dos parámetros se utiliza el método de los mínimos cuadrados, que consiste en ajustar estos parámetros para minimizar la suma del cuadrado de los residuales:

$$S = \sum_{i=1}^n (y_i - a - bx_i)^2 \quad (7)$$

Para poder aplicar este método se debe considerar que no existe error en la variable  $x$ , siendo este considerado sólo en la variable  $y$  por mediación de la varianza de repetición pura (obtenida a partir de replicados). Además, dicha varianza debe permanecer constante en los distintos niveles (homocedasticidad) y los valores de  $y$  deben estar normalmente distribuidos en cada nivel. Esto se puede observar en la gráfica de residuales.



Las ecuaciones para calcular la pendiente y ordenada son:

$$b = \frac{\sum_i^n \{(x_i - \bar{x})(y_i - \bar{y})\}}{\sum_i^n (x_i - \bar{x})^2} \quad (8)$$

$$a = \bar{y} - b\bar{x} \quad (9)$$

Los errores de estos parámetros se calculan a partir de la varianza de regresión, que estima los errores aleatorios en la dirección y.

$$s_{y/x} = \sqrt{\frac{\sum_i^n (y_i - \hat{y}_i)^2}{n-2}} \quad (10)$$

Donde

$$\hat{y}_i = a + bx_i \quad (11)$$

A partir de esta se tiene la desviación estándar de la pendiente y de la ordenada:

$$s_b = \frac{s_{y/x}}{\sqrt{\sum_i^n (x_i - \bar{x})^2}} \quad (12)$$

$$s_a = s_{y/x} \sqrt{\frac{\sum_i^n x_i^2}{\sum_i^n (x_i - \bar{x})^2}} \quad (13)$$

El último parámetro que vamos a considerar es el coeficiente de correlación,  $r$ , que da cuenta de la bondad del ajuste de la curva a los datos experimentales. Su valor oscila entre 0 y 1, valores de  $r$  próximos a 1 implican un mejor ajuste.

$$r = \frac{\sum_i^n \{(x_i - \bar{x})(y_i - \bar{y})\}}{\sqrt{\left[ \sum_i^n (x_i - \bar{x})^2 \right] \left[ \sum_i^n (y_i - \bar{y})^2 \right]}} \quad (14)$$

Si tenemos una muestra de señal  $y_0$ , se puede calcular la concentración correspondiente a la misma como:

$$x_0 = \frac{y_0 - a}{b} \quad (15)$$

Y el error de esta concentración viene dado por:

$$s_{x_0} = \frac{s_{y/x}}{b} \sqrt{1 + \frac{1}{n} + \frac{(y_0 - \bar{y})^2}{b^2 \sum_i^n (x_i - \bar{x})^2}} \quad (16)$$

### 1.8. Algunas fórmulas en EXCEL.

Parámetro o función	Fórmula
Media	=PROMEDIO()
Desviación estándar	=DESVEST()
Varianza	=VAR()
Raíz cuadrada	=RAIZ()
Número de datos	=CONTAR()
Mediana	=MEDIANA()
Pendiente	=PENDIENTE()
Ordenada en el origen	=INTERSECCION.EJE()
Coefficiente de correlación	=COEF.DE.CORREL()
Estimación lineal	=ESTIMACION.LINEAL()

La última es muy útil cuando se quieren obtener los errores asociados a la pendiente y la ordenada en el origen. Se explicará mas adelante.

## **2. Validación de métodos analíticos.**

Un método analítico es un procedimiento que obtiene unos resultados a partir del análisis químico de muestras. Antes de ser usado, es necesario comprobar que dicho método cumple unos determinados requisitos que dependerán de la aplicación que se le quiere dar. Este proceso de verificación se conoce como validación y debe establecer las características de funcionamiento y las limitaciones del método, así como la identificar las influencias que pueden cambiar dichas características [1]. A continuación se detallan los distintos parámetros relacionados con la calidad de un método analítico que deben ser objeto de estudio en el proceso de validación.

### *2.1. Selectividad y especificidad.*

En todo método analítico hay un paso fundamental en el que se produce la medida de una señal relacionada con la concentración (análisis cuantitativo) o la presencia (análisis cualitativo) de un determinado analito. A veces, la presencia de otros compuestos, química o físicamente similares al analito, pueden contribuir a dicha señal, pudiendo causar un error sistemático en la determinación del analito. Estos compuestos pasan a denominarse interferencias. Es por tanto necesario asegurarse de donde procede la señal que se está midiendo. De aquí surge en concepto de selectividad, entendiéndose esta como el parámetro que define de forma cualitativa la extensión en que otras sustancias interfieren en la determinación de un analito usando un método determinado. Al máximo de selectividad se le denomina especificidad, es decir, un método será específico cuando la señal medida solo proviene del analito de interés [2]. Un estudio de selectividad implica investigar la capacidad del método para medir al analito en presencia de interferencias añadidas a la muestra [1], no siendo esto objeto del presente texto.

Hasta ahora, solo hemos considerado las interferencias que pueden producir un error sistemático constante, es decir, sustancias que contribuyen a la señal medida con un valor determinado. Pero existen otro tipo de interferencias, relacionadas con efectos inespecíficos, que pueden producir errores proporcionales. Estas son las designadas como efecto matriz y, aunque no existe ninguna aplicación en Excel específica para cuantificar este efecto, hablaremos de el más adelante.

### *2.2. Exactitud.*



La exactitud da cuenta de la proximidad entre el resultado obtenido por un método y el valor “real”. Como se trata de un parámetro cualitativo (un método será muy o poco exacto) para describir el error asociado a un resultado, hace falta un parámetro que permita medir la exactitud cuantitativamente. Este parámetro es el error que, según la guía para validación de métodos de la EURACHEM [1], se divide en dos componentes, la veracidad y la precisión. La veracidad es un parámetro que expresa la proximidad de la media de una serie de resultados obtenidos con el método al valor real. Generalmente se expresa en términos de error o bias, definido como la diferencia entre el resultado de medida y el valor real [2]. La precisión da cuenta de la proximidad entre sí de esta serie de resultados. Si expresamos la exactitud en función del error  $E_i$ , para una determinada medida  $i$ :

$$E_i = x_i - \mu = (x_i - \bar{x}) + (\bar{x} - \mu) \quad (17)$$

En el caso de tener un replicado consistente en  $n$  medidas, el error cuadrático se calcularía como:

$$\sum_{i=1}^{i=n} E_i^2 = \sum_{i=1}^{i=n} (x_i - \bar{x})^2 + n(\bar{x} - \mu)^2 \quad (18)$$

Si dividimos por  $n$ , obtenemos el error cuadrático medio:

$$\langle E^2 \rangle = \frac{\sum_{i=1}^{i=n} (x_i - \bar{x})^2}{n} + (\bar{x} - \mu)^2 \quad (19)$$

Si  $n$  es suficientemente grande, el primer término tiende a la varianza de la medida (dividiendo entre  $n-1$  en vez de  $n$ ) y podría expresarse:

$$\langle E^2 \rangle = s^2 + (\bar{x} - \mu)^2 \quad (20)$$

Con lo cual vemos como el error cuadrático medio correspondiente a la exactitud se equipara a la suma de la varianza de repetición pura (error aleatorio) más el cuadrado de la veracidad (error sistemático).

La exactitud es máxima cuando el sesgo es nulo y la dispersión aleatoria es mínima.

### 2.3. Veracidad.

Como se ha dicho anteriormente, la veracidad da cuenta de lo cercano del resultado medio obtenido con el método y el valor real. Para asegurar la veracidad es necesario comparar el valor medio medido con un valor de referencia (que será el valor real o el considerado como tal). Esto se realiza, generalmente comparando los resultados obtenidos por el método con el valor de un material de referencia certificado o comparando con un método ya validado (que suele haber sido sometido a un ejercicio de intercomparación). En caso de no existir materiales de referencia o no tener posibilidad de aplicar el método de referencia, se puede simular la matriz de la muestra añadiendo al analito en una cantidad conocida y efectuar un ensayo de recuperación. Como última instancia, se pueden realizar dichos ensayos de recuperación añadiendo al analito sobre la propia muestra [3].

Como se explicará más adelante, para comparar dos valores será necesario conocer la precisión con la que este ha sido obtenido, por ello la exactitud se evalúa a partir de la veracidad en conjunto con la precisión.

### 2.4. Precisión.

Cuando realizamos una medida se deben obtener replicados, a partir de los cuales se puede observar el grado de concordancia que estos resultados tienen entre sí. Como se dijo anteriormente, este grado de concordancia se denomina precisión. La precisión se divide en dos parámetros repetitividad y reproducibilidad.

#### 2.4.1. Repetitividad.

Se define la repetitividad como la desviación estándar obtenida al analizar una misma muestra varias veces, en un periodo de tiempo corto, sin cambiar de equipo de medida, reactivos o analista.

#### 2.4.2. Reproducibilidad.

La reproducibilidad se define como la desviación estándar obtenida al analizar varias veces la muestra en días distintos, pudiendo variar condiciones tales como el equipo, reactivos o analistas. Se habla de reproducibilidad interlaboratorio cuando las medidas se realizan en laboratorios distintos (ensayos de intercomparación). La reproducibilidad intralaboratorio se realiza en un solo laboratorio, preparando los patrones de calibración cada vez, en distintos días y cambiando en la medida de lo posible de equipo y analista. A esta reproducibilidad intralaboratorio se la conoce como precisión intermedia.

### 2.5. Sensibilidad.

Según la IUPAC la sensibilidad se define como el cociente entre la señal medida y la concentración de analito. Cuando trabajamos con curvas de calibrado, la sensibilidad es la pendiente de la curva a una concentración dada. En caso de ser una recta, coincide con la pendiente de la misma.

### 2.6. Límites de detección y cuantificación.

Estos parámetros están relacionados con la sensibilidad del método. El límite de detección (LOD) se define como la cantidad de analito presente que se deriva de la menor señal analítica que puede detectarse con certeza razonable ( $Y_{LOD}$ ) [4]. Si analizamos un blanco de la muestra varias veces y obtuviéramos la señal media ( $\bar{Y}_0$ ) y su desviación estándar,  $s_{\bar{Y}_0}$ , las señales mayores a  $Y_{LOD} = \bar{Y}_0 \pm 3s_{\bar{Y}_0}$  caerían fuera de la gaussiana del blanco analítico y podríamos asignarlas a la presencia de analito. Si tenemos una recta de calibrado de ecuación  $y = bx + a$ , el LOD se calcularía como:

$$X_{LOD} = \frac{Y_{LOD} - a}{b} \quad (21)$$

Como la señal del blanco debe coincidir con la ordenada del blanco analítico (puesto que es la señal que se obtendría a una concentración de analito igual a cero), el LOD se calculará como:

$$X_{LOD} = \frac{3s_a}{b} \quad (22)$$

siendo  $s_a$  la desviación estándar de la ordenada.

El límite de cuantificación (LOQ) se define como la cantidad de analito presente derivada de la menor señal analítica que puede ser medida con precisión y exactitud. De manera análoga al LOD, la IUPAC propone como LOQ al valor  $Y_{LOQ} = \bar{Y}_0 \pm 10s_{\bar{Y}_0}$ , que lleva a una concentración, a partir de la recta de calibrado:

$$X_{LOQ} = \frac{10s_a}{b} \quad (23)$$

### 2.7. Linealidad e intervalo dinámico

La linealidad a veces se examina a partir del coeficiente de correlación, considerando lineal el calibrado si  $r > 0.999$ . Otro método se basa en que la desviación estándar

relativa de la pendiente no supere el 5%. Esta es la conocida como linealidad online. El porcentaje de linealidad se calcula como [5]:

$$\%Lin = 100 \left( 1 - \frac{s_b}{b} \right) \quad (24)$$

Estos parámetros no son del todo fidedignos, siendo lo correcto realizar un análisis estadístico de los residuales mediante el análisis de la varianza (ANOVA) o la aproximación práctica de Huber [6]. Ambos procedimientos los explicaremos detenidamente cuando estudiemos la aplicación de EXCEL.

El intervalo dinámico o rango lineal es el rango de concentraciones donde existe una relación lineal entre estas y las señales analíticas. En este intervalo, los valores deben presentar un nivel aceptable de precisión y exactitud, con lo cual, no pueden ser inferiores al límite de cuantificación.

### 2.8. Robustez

Es la inercia que presenta un método analítico a modificar su señal cuando tienen lugar pequeños cambios en las condiciones ambientales u operativas, que se consideran las variables que gobiernan el experimento de medida. Para el estudio de robustez, estas condiciones se modifican ligeramente, siguiendo un determinado diseño experimental, y se comprueba su influencia en la señal analítica. Con ello, se puede concluir que variables son más significativas a la hora de realizar las medidas y por tanto, deben ser mejor controladas.

### 2.9. Ámbito de validación de métodos analíticos

Como ya hemos dicho, el tipo de validación que debemos realizar va a depender de los requisitos de nuestro método.

Parámetro	Análisis Cualitativo	Análisis Cuantitativo	Análisis de trazas
Especificidad	X	X	X
Precisión		X	X
Veracidad		X	X
Linealidad		X	X
Rango			X
LOD/LOQ	LOD		LOQ
Robustez	X	X	X

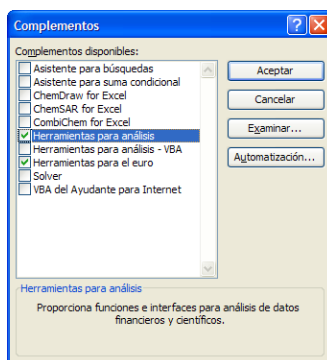
### 3. EXCEL como herramienta en el proceso de validación.

Las utilidades estadísticas de EXCEL que emplearemos en validación serán:

- Prueba F de Snedecor-Fisher
- Prueba t de Student
- ANOVA de un factor
- Estadística descriptiva (media, mediana, varianza...)
- Regresión

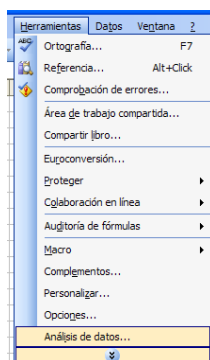
#### 3.1. Herramientas para análisis.

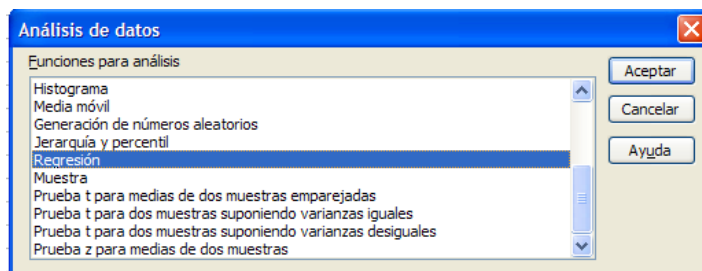
Aunque todo esto se puede hacer de forma manual, es decir, introduciendo formulas en la hoja de cálculo, EXCEL incluye un paquete de herramientas para realizar este tipo de análisis de manera sencilla. El primer paso será comprobar que está instalado o activado. Para ello hay que abrir el menú Herramientas>Complementos y ver si aparece la opción Herramientas para análisis.



Si no aparece debemos activar el paquete en C:\Archivos de programa\Microsoft Office\OFFICE11\Macros\Análisis\ANALYS32.

En este momento dispondremos de la herramienta análisis de datos.





### 3.2. Prueba F de Snedecor- Fisher.

En una prueba estadística empleada para la comparación de varianzas.

Sean dos conjuntos de datos:

$\{x_1\}$  con varianza  $s_1^2$  y tamaño  $n_1$

$\{x_2\}$  con varianza  $s_2^2$  y tamaño  $n_2$

Consideramos la hipótesis nula que dice que las varianzas no presentan diferencias

significativas entre sí.  $H_0: \sigma_1^2 = \sigma_2^2$ . Si  $s_2^2 \geq s_1^2$ ,  $F = \frac{s_2^2}{s_1^2} \geq 1$ . Se compara con  $F_{\text{tab}}(0.05,$

$n_2-1, n_1-1)$ . Si  $F < F_{\text{tab}}$  admitimos  $H_0$ . Los valores de F para una cola se pueden obtener usando la función =DISTR.F.INV(p;n<sub>2</sub>-1;n<sub>1</sub>-1), siendo p la probabilidad (normalmente 0.05, es decir 95% de nivel de confianza)

#### 3.2.1. Aplicaciones

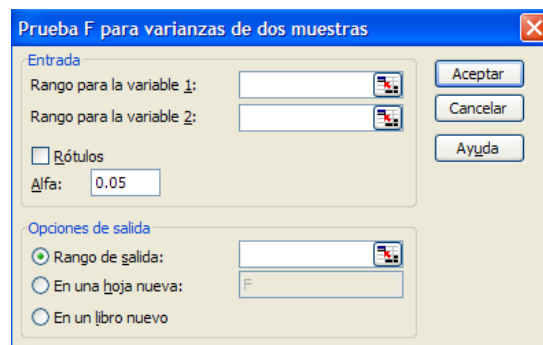
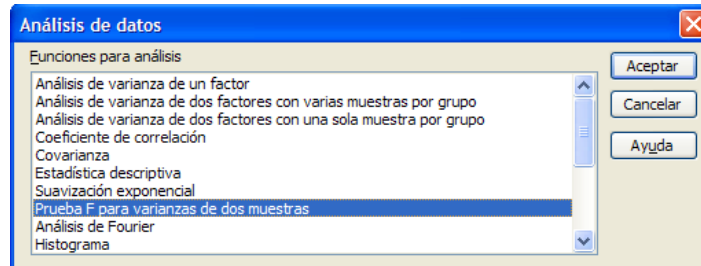
- Comprobar que dos métodos son igual de precisos.
- Comprobar que las varianzas son iguales para ver que tipo de test de comparación de medias usar posteriormente.
- Comprobar que en la reproducibilidad no interviene factores distintos a los encontrados en el estudio de repetitividad.

#### 3.2.2. Resolución con EXCEL

Supongamos que hemos realizado una serie de medidas de una muestra con dos métodos analíticos y queremos comprobar si ambos son igual de precisos. Lo primero es tener los datos en EXCEL.

	A	B
1	Método A	Metodo B
2	2	2.3
3	2	2.5
4	2.2	2.6
5	2.3	2.6
6	2.5	2.6
7	2.6	2.6
8		2.4
9		2.3

Del menú Herramientas>Análisis de datos usamos la herramienta “Prueba F para varianza de dos muestras”



Pulsamos en la flecha roja del cuadro rango para la variable 1 y seleccionamos el rango A2:A7. En rango para la variable 2 seleccionamos el rango B2:B9. Si marcamos Rótulos, el programa considera que en la primera celda del rango seleccionado se encuentra el nombre del conjunto de datos (deberíamos haber seleccionado A1:A7 y B1:B9). Alfa es el indica la probabilidad o nivel de confianza que usaremos (0.05 es un 95%). En rango de salida se indica la primera celda del rango donde se colocarán los resultados. Es aconsejable que sea en un extremo para que no sobrescriba nada. Otra opción es hacerlo en una hoja nueva. Supongamos que hemos elegido la celda D1 como rango de salida. El resultado será:

	D	E	F
Prueba F para varianzas de dos muestras			
		Método A	Método B
Media		2.26666667	2.4875
Varianza		0.06266667	0.01839286
Observaciones		6	8
Grados de libertad		5	7
F		3.40711974	
P(F<=f) una cola		0.0704931	
Valor crítico para F (una cola)		3.97152315	

Se comprueba que  $F < F_{crit}$ , ambos métodos son igual de precisos.

Si necesitásemos el valor de F crítico para comparar las varianzas a manos se puede obtener de tablas o bien con la formula =DISTR.F.INV(0.05,5,7), en este caso.

### 3.3. Prueba t de Student.

#### 3.3.1. Comparación de valores medios de dos conjuntos de datos.

Dados dos conjuntos de datos:

Conjunto	Media	Varianza	Número de datos
1	$\bar{x}_1$	$s_1^2$	$n_1$
2	$\bar{x}_2$	$s_2^2$	$n_2$

Para decidir si las medias muestrales  $\bar{x}_1$  y  $\bar{x}_2$  difieren significativamente se toma la hipótesis nula  $H_0: \mu_1 = \mu_2$ .

Primero se comprueba que  $s_1^2 = s_2^2$  mediante la prueba F de Fisher. Así tenemos dos posibilidades:

a) Si  $s_1^2 = s_2^2$  se calcula una varianza promedio

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \quad (25)$$

Se calcula el estadístico t:

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad (26)$$

, que tendrá  $n_1 + n_2 - 2$  grados de libertad. Este valor de t se compara con el valor tabulado  $t(0.05, n_1 + n_2 - 2)$ . Si  $t < t_{crit}$  se acepta la hipótesis nula, es decir, las dos medias son estadísticamente iguales.



b) Si  $s_1^2 \neq s_2^2$  se calcula el estadístico t como:

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (27)$$

Ahora los grados de libertad se deben calcular como:

$$v_{eff} = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\left(\frac{s_1^4}{n_1^2(n_1-1)} + \frac{s_2^4}{n_2^2(n_2-1)}\right)} \quad (28)$$

El valor de t se compara con  $t(0.05, v_{eff})$ .

Los valores de t se pueden obtener con la función =DISTR.T.INV(p;g.l.). Estos valores son adecuados para contraste de dos colas, para una cola se toma probabilidad 2\*p (siendo p el nivel de confianza deseado).

### 3.3.1.1. Aplicaciones.

- Comparación del valor medio obtenido con para un material de referencia aplicando un determinado método y el valor de un certificado.
- Comparación de los resultados obtenidos con un método en desarrollo y uno de referencia.

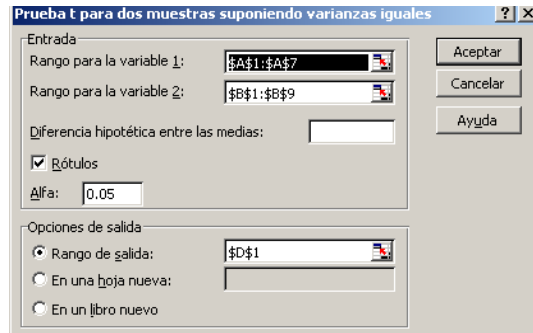
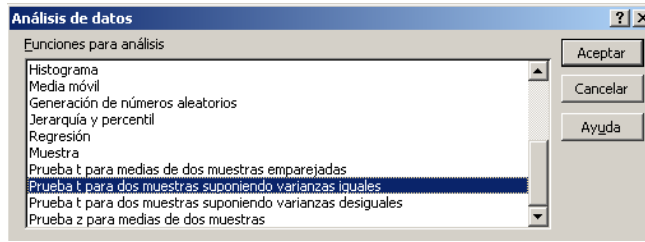
### 3.3.1.2. Resolución con EXCEL.

a) Conjunto de datos con varianzas estadísticamente similares.

Supongamos que partimos del ejemplo anterior, en el que mediante el test de Fisher se dedujo que las varianzas eran similares.

	A	B
1	Método A	Metodo B
2	2	2.3
3	2	2.5
4	2.2	2.6
5	2.3	2.6
6	2.5	2.6
7	2.6	2.6
8		2.4
9		2.3

En Herramientas>Análisis de datos se escoge la opción “Prueba t para dos muestras suponiendo varianzas iguales”



La forma de escoger los rangos, nivel de confianza y rango de salida son similares a la prueba F. Nótese que es posible tomar distintas hipótesis nulas en la casilla Diferencia hipotética entre las medias. El resultado es:

D	E	F	G
Prueba t para dos muestras suponiendo varianzas iguales			
	<i>Método A</i>	<i>Metodo B</i>	
Media	2.26666667	2.4875	
Varianza	0.06266667	0.01839286	
Observacione	6	8	
Varianza agru	0.03684028		
Diferencia hip	0		
Grados de lib	12		
Estadístico t	-2.13039327		
P(T<=t) una c	0.02725833		
Valor crítico c	1.78228674		
P(T<=t) dos c	0.05451667		
Valor crítico c	2.17881279		

Como en valor absoluto el t calculado es menor que el t crítico (contraste de dos colas), la hipótesis nula se toma como cierta y las dos medias muestrales no difieren significativamente. El hecho de usar un contraste de dos colas implica que no nos interesa el sentido en que difieren las medias (es decir si la diferencia es positiva o negativa). Si pretendemos ver solo si un valor aumenta en relación al otro se usa un contraste de una cola.

b) conjunto de datos con varianzas distintas.

Considérese el conjunto de datos:

	A	B
1	Método A	Metodo B
2	1	2.3
3	1	2.5
4	2.2	2.6
5	2.3	2.6
6	1.5	2.6
7	2.6	2.6
8		2.4
9		2.3

Al aplicar el test F:

	D	E	F
Prueba F para varianzas de dos muestras			
		<i>Método A</i>	<i>Metodo B</i>
Media		1.76666667	2.4875
Varianza		0.48266667	0.01839286
Observacione		6	8
Grados de lib		5	7
F		26.2420712	
P(F<=f) una c		0.00021364	
Valor crítico p		3.97152178	

Las varianzas difieren significativamente. A la hora de realizar el contraste t, se debe elegir “Prueba t para dos muestras suponiendo varianzas desiguales”. Los rangos, probabilidad e hipótesis nula se introducen de igual modo que en el ejemplo anterior. El resultado es:

Prueba t para dos muestras suponiendo varianzas desiguales			
	<i>Método A</i>	<i>Metodo B</i>	
Media	1.76666667	2.4875	
Varianza	0.48266667	0.01839286	
Observacione	6	8	
Diferencia hip	0		
Grados de lib	5		
Estadístico t	-2.50592335		
P(T<=t) una c	0.02704842		
Valor crítico c	2.01504918		
P(T<=t) dos c	0.05409685		
Valor crítico c	2.57057764		

Comparando con el valor de t crítico para dos colas, las medias no difieren significativamente. Nótese que los grados de libertad efectivos son 5 (en realidad 5.2868, pero se ajusta al entero más próximo).

### 3.3.2. Prueba t para valores emparejados.

Supongamos que tenemos un conjunto de muestras de valores  $x_1, \dots, x_i, \dots, x_n$ . Nótese que se trata de n muestras distintas, por lo que las cantidades  $x_1 \neq \dots \neq x_i \neq \dots \neq x_n$ . Se determinan los valores mediante dos métodos y se pretende comprobar que no existen diferencias significativas entre los resultados obtenidos por ambos. Al aplicar cada método se obtienen los valores  $x_{11}, \dots, x_{1i}, \dots, x_{1n}$  con el método 1 y  $x_{21}, \dots, x_{2i}, \dots, x_{2n}$  con el

método 2. Se calculan las diferencias  $d_i = x_{1i} - x_{2i}$ , el valor medio de todas las diferencias y su desviación estándar:

$$\bar{d} = \frac{\sum_i^n d_i}{n} \quad (29)$$

$$s_d = \sqrt{\frac{\sum_i^n (d_i - \bar{d})^2}{n-1}} \quad (30)$$

Si los métodos son equivalentes y llevan al mismo resultado las diferencias  $d_i$  deben ser nulas. Esto es lo mismo que considerar la hipótesis nula  $H_0: \mu_d = 0$ . Se calcula el estadístico t:

$$t = \bar{d}\sqrt{n} / s_d \quad (31)$$

El número de grados de libertad de t es n-1. Se compara este valor con t(0.05,n-1).

### 3.3.2.1. Aplicaciones.

Se aplica en el caso de que las muestras analizadas por los dos métodos no sean grandes y sea necesario utilizar distintas muestras en el estudio. Es decir, se analizan distintas muestras por ambos métodos y no una misma muestra replicada por ambos métodos.

### 3.3.2.2. Resolución con EXCEL

Consideramos el ejemplo de la determinación de paracetamol en diez pastillas de distintos lotes mediante un método espectrofotométrico y otro de espectroscopía IR [7].

	A	B
1	Método A	Método B
2	84.63	83.15
3	84.68	83.72
4	84.08	83.84
5	84.41	84.2
6	83.82	83.92
7	83.55	84.16
8	83.92	84.02
9	93.69	83.6
10	84.06	84.13
11	84.03	84.24

Se utiliza la “Prueba t para dos muestras emparejadas” del menú Herramientas>Análisis de datos. El formulario de entrada de datos es similar a los de los ejemplos anteriores. El resultado es:

	D	E	F	G
Prueba t para medias de dos muestras emparejadas				
		Método A	Método B	
Media		84.057	83.898	
Varianza		0.11391222	0.11437333	
Observaciones		10	10	
Coficiente de correlación		-0.42321351		
Diferencia hipotética		0		
Grados de libertad		9		
Estadístico t		0.88211054		
P(T<=t) una cola		0.20033133		
Valor crítico de t una cola		1.83311386		
P(T<=t) dos colas		0.40066266		
Valor crítico de t dos colas		2.26215889		

Como el valor experimental (0.88) es menor que el crítico (2.26) los métodos no proporcionan resultados significativamente diferentes.

### 3.4. ANOVA de un factor.

El análisis de la varianza (ANOVA) es un método que se emplea para comparar resultados obtenidos por distintos métodos, laboratorios, analistas, etc, cuando el número de medias obtenidas es superior a dos. También permite separar las contribuciones de uno o más factores a la varianza global del sistema.

En el caso de realiza medidas replicadas variando uno de los factores anteriormente mencionados siempre existen dos fuentes de error, la primera, el error aleatorio de medida, la segunda los errores debidos al cambio de método, laboratorio, analista, etc. Mediante el ANOVA se puede controlar el error introducido por esta segunda fuente, con lo que se habla de ANOVA de un factor. Si tuviésemos un caso en que el análisis se realiza por distintos laboratorios utilizando distintos métodos, deberíamos controlar dos factores y se utilizaría un ANOVA de dos factores.

Consideremos que realizamos un mismo análisis por  $i$  métodos distintos ( $i$  varía de 1 a  $h$ ) y realizamos  $j$  replicados ( $j$  varia de 1 a  $n$ ) con cada método.

La media obtenida para cada método será:

$$\bar{x}_i = \frac{\sum_{j=1}^n x_{ij}}{n} \quad (32)$$

La media global obtenida por todos los métodos se calculará como:

$$\bar{x} = \frac{\sum_{i=1}^h \bar{x}_i}{h} \quad (33)$$

							Media
Método 1	$x_{11}$	$x_{12}$	...	$x_{1j}$	...	$x_{1n}$	$\bar{x}_1$
Método 2	$x_{21}$	$x_{22}$	...	$x_{2j}$	...	$x_{2n}$	$\bar{x}_2$
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
Método i	$x_{i1}$	$x_{i2}$	...	$x_{ij}$	...	$x_{in}$	$\bar{x}_i$
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
Método h	$x_{h1}$	$x_{h2}$	...	$x_{hj}$	...	$x_{hn}$	$\bar{x}_h$
Media global							$\bar{x}$

Para cada método se tiene una varianza calculada como:

$$s_i^2 = \frac{\sum_j^n (x_{ij} - \bar{x}_i)^2}{n-1} \quad (34)$$

En el ANOVA la hipótesis nula es que todas las muestras se extraen de una población de media  $\mu$  y varianza  $\sigma^2$ . Esta varianza se puede estimar de dos formas, estudiando la variación dentro de las muestras y la variación entre muestras.

La varianza debida al error puramente aleatorio,  $s_{PE}^2$ , también conocida como varianza dentro de las muestras (o del grupo), se calcula como promedio de las varianzas individuales  $s_i^2$ :

$$\sigma^2 \approx s_{PE}^2 = \frac{\sum_i^h s_i^2}{h} = \frac{\sum_i^h \sum_j^n (x_{ij} - \bar{x}_i)^2}{h(n-1)} \quad (35)$$

La varianza debida al factor (en nuestro caso a los distintos métodos usados), conocida como varianza entre muestras (o entre grupos), es una estimación de la varianza de la población,  $\sigma^2$ . Si todas las medidas se extraen de una población de varianza  $\sigma^2$ , entonces sus medias (obtenidas variando el método empleado, es decir, variando un factor controlado) proceden de una población de varianza  $\sigma^2/n$  (Véase sección 1.4). En general, la varianza debida al factor se calcula como:

$$\sigma^2 \approx s_{Factor}^2 = \frac{n \sum_i^h (\bar{x}_i - \bar{x})^2}{h-1} \quad (36)$$

Si la hipótesis nula es correcta, las dos estimaciones no deberían diferir significativamente, es decir, que la variación debida al factor no debe ser significativamente distinta a la varianza de error puro. Esto es lo mismo que decir que el hecho de variar el factor (el método de medida en nuestro ejemplo) no introduce un error significativo en comparación con el puramente aleatorio. Se comprueba esto mediante un test de Fisher, calculando F como:

$$F = \frac{s_{Factor}^2}{s_{PE}^2} \quad (37)$$

Se compara F con el valor tabulado F(0.05, h-1, h(n-1)) para contraste de una cola (puesto que nos interesa saber si la varianza del factor es mayor que la de error puro, no solo si difieren). Si el valor de F calculado es menor que el tabulado, no existen diferencias significativas entre las dos varianzas, es decir, el hecho de variar el factor (cambiar de método) no introduce un error significativo.

Si la hipótesis nula es verdadera existe una tercera forma de estimar la varianza que consiste en tratar los datos como si fueran una muestra grande de varianza:

$$s_{Total}^2 = \frac{\sum_i^h \sum_j^n (x_{ij} - \bar{x})^2}{hn-1} \quad (38)$$

Pero esta fórmula no se usa en el contraste.

Hasta ahora hemos supuesto que la hipótesis nula es verdadera. Pero puede que no lo sea, es decir, que F sea mayor que el valor de F crítico. Un resultado significativo en el ANOVA puede deberse a varias cosas: que una media difiera mucho a las demás, que todas las medias difieran entre sí o que existan varios grupos de medias distintas, etc. Para poder ver lo que ocurre se colocan las medias ordenadas de forma creciente y se comparan las diferencias de los valores adyacentes con una cantidad denominada la mínima diferencia significativa. Ésta se calcula como:

$$MDS = s_{PE} \sqrt{\frac{2}{n}} \cdot t_{h(n-1)} \quad (39)$$

Otra posibilidad es comparar las medias dos a dos mediante el test de de la *genuina diferencia significativa de Tukey*. Si una vez realizados los cálculos de ANOVA se han encontrado diferencias significativas entre la varianza de un factor y la del error puro, debemos averiguar que nivel de ese factor es el responsable de los errores sistemáticos que producen dichas diferencias. Con este fin se realiza el ensayo Tukey-HSD (Honest Significant Difference) [8], por el que se comparan dos a dos las medias obtenidas a cada nivel del factor. Aquel nivel que presente un valor promedio significativamente diferente del resto será el responsable del sesgo detectado y se elimina para recalcular el ANOVA.

Supongamos que queremos comparar dos niveles  $p$  y  $q$  del factor (método en nuestro caso). Para que no haya diferencia significativa entre los valores promedio de estos dos niveles,  $\bar{x}_p$  y  $\bar{x}_q$ , se debe demostrar que ambos se corresponden con las medias muestrales de una misma población, de distribución normal, de media  $\mu$  y varianza  $\sigma^2$ . De este modo, si  $\bar{x}_p$  es la media muestral de una población de media  $\mu_p$  y  $\bar{x}_q$  es la media muestral de una población de media  $\mu_q$ , se debe cumplir, como hipótesis nula, que  $\mu_p = \mu_q$ .

Se define el intervalo de confianza de Tukey como:

$$\mu_p - \mu_q \in (\bar{x}_p - \bar{x}_q) \pm \frac{q_{r,v}^\alpha}{\sqrt{2}} \sqrt{s_{PE}^2 \left( \frac{1}{n_p} + \frac{1}{n_q} \right)} \quad (40)$$

Donde,  $q_{r,v}^\alpha$  es el valor tabulado de la distribución del rango estandarizado [9], que depende del numero de repeticiones,  $r$ , y de los grados de libertad utilizados en la desviación estándar que se usa en el calculo de este intervalo,  $v$ , así como del nivel de confianza,  $\alpha$ .

Teniendo en cuenta que, la varianza utilizada en la expresión anterior es una estimación de la de la población y debe considerar todas las muestras, se sustituye ésta por la varianza del error puro. Además, si en nuestro caso  $n_p = n_q = n$ , el intervalo de confianza de Tukey queda:



$$\mu_p - \mu_q \in (\bar{x}_p - \bar{x}_q) \pm \frac{q_{r,v}^\alpha}{\sqrt{2}} s_{PE} \sqrt{\frac{1}{n}} \quad (41)$$

Como  $\mu_p - \mu_q = 0$ , el intervalo de confianza de Tukey debe contener el cero, si no es así, se rechaza la hipótesis nula.

$$(\bar{x}_p - \bar{x}_q) - \frac{q_{r,v}^\alpha}{\sqrt{2}} s_{PE} \sqrt{\frac{1}{n}} < 0 < (\bar{x}_p - \bar{x}_q) + \frac{q_{r,v}^\alpha}{\sqrt{2}} s_{PE} \sqrt{\frac{1}{n}} \quad (42)$$

En la práctica, se parte de la hipótesis nula,  $\mu_p = \mu_q$ , y se calcula un valor de q experimental según:

$$q_{\text{exp}} = \frac{\bar{x}_p - \bar{x}_q}{s_{PE} \sqrt{\frac{1}{n}}} \quad (43)$$

Se compara este valor experimental de q con el valor tabulado [9]. Si  $q_{\text{exp}} > q_{\text{tab}}$  se rechazaría la hipótesis nula, con lo que las dos medias presentarían diferencias significativas entre sí. El nivel que presente diferencias significativas con el resto, se considera como causa del efecto significativo en la varianza total. Una tabla de valores de la distribución del rango estandarizado se incluye en el ANEXO 1.

#### 3.4.1. Aplicaciones.

El análisis de la varianza se aplica, como se dijo anteriormente, en la comparación de varios métodos, laboratorios, analistas. etc.

Es muy útil para detectar resultados anómalos en los ensayos de intercomparación empleados para certificar materiales de referencia o desarrollar métodos de referencia.

También se utiliza para calcular la precisión intermedia de un método cuando se analizan muestras de referencia en distintos días. Este ejemplo lo explicaremos más adelante.

Otra aplicación tiene que ver con la regresión lineal, puesto que EXCEL realiza un análisis de la varianza de los residuales. Esto lo veremos en el apartado correspondiente a la estimación lineal.

Ensayos de recuperación.

## 3.4.2. Resolución con EXCEL.

Supongamos que realizamos medidas de fluorescencia a muestras de quinina almacenadas de cuatro formas distintas (ejemplo adaptado de [7]): A) disoluciones recién preparadas, B) almacenada en oscuridad, C) con luz tenue y D) luz brillante. Los datos los ordenamos como se observa a continuación:

	A	B	C	D
1	Método A	Método B	Método C	Método D
2	102	101	97	90
3	100	101	95	92
4	101	104	99	94
5				

En el menú Herramientas>Análisis de datos seleccionamos “Análisis de varianza de un factor”. El formulario de entradas de datos es el siguiente:

Como el rango de entrada seleccionamos A1:D4, señalando la casilla de rótulos para que tome de la primera fila el nombre de los grupos. Hay que cerciorarse que tenemos marcado que los datos se agrupan por columnas. El rango de salida funciona como anteriormente se ha descrito. El resultado final es el siguiente:

	F	G	H	I	J	K	L
1	Análisis de varianza de un factor						
2							
3	<b>RESUMEN</b>						
4	<i>Grupos</i>	<i>Cuenta</i>	<i>Suma</i>	<i>Promedio</i>	<i>Varianza</i>		
5	Método A	3	303	101	1		
6	Método B	3	306	102	3		
7	Método C	3	291	97	4		
8	Método D	3	276	92	4		
9							
10	<b>ANÁLISIS DE VARIANZA</b>						
12	<i>Origen de las variaciones</i>	<i>Suma de cuadrados</i>	<i>Grados de libertad</i>	<i>Promedio de los cuadrados</i>	<i>F</i>	<i>Probabilidad</i>	<i>Valor crítico para F</i>
13	Entre grupos	186	3	62	20.6666667	0.000400152	4.066180281
14	Dentro de los grupos	24	8	3			
15							
16	Total	210	11				
17							

La columna de promedio de cuadrados indica el valor de varianza de error puro (dentro de los grupos) y la debida al cambio en el factor (entre grupos), que en este caso es el cambio en condiciones de almacenamiento. Como se observa el valor de F calculado (20.66) es mayor que el crítico para un nivel de confianza del 95 % y 3 y 8 grados de

libertad (4.066). Existen entonces diferencias significativas entre las medias obtenidas tras las distintas formas de almacenamiento de las muestras. Los test de mínima diferencia significativa o de Tukey deben realizarse manualmente (EXCEL no implementa los mismos).

### 3.4.3. La precisión a partir del análisis de la varianza.

Hasta ahora se ha aplicado el ANOVA para la comparación de múltiples medias. En este caso se va a emplear para separar y estimar distintas fuentes de variación. Supongamos que realizamos un estudio de reproducibilidad a un solo nivel de concentración. Tendremos dos fuentes de variación, la debida al error puramente aleatorio, es decir, el error de replicación, y el debido al cambio de condiciones (día, reactivos, analista, laboratorio...) que implica el hecho de estudiar la reproducibilidad. Lo ideal sería añadir otro factor, realizando el estudio en distintos niveles de concentración, pero entonces realizaríamos un ANOVA de dos factores. Es más, al prevalecer el factor de las condiciones de trabajo sobre los niveles de concentración deberíamos realizar un ANOVA de dos factores anidados.

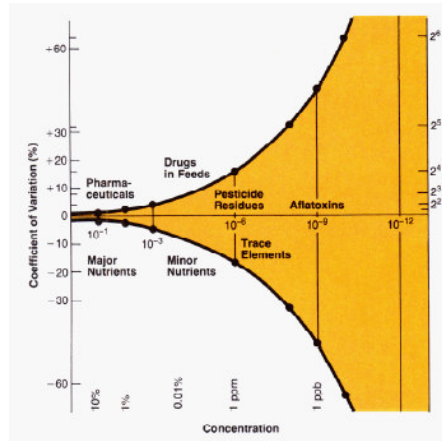
Consideremos que realizamos  $n$  replicados durante  $h$  días (reproducibilidad entre días, minimalista). La varianza de repetitividad o de error puro vendrá dada por la ecuación 35, mientras que la debida al factor vendrá dada por la ecuación 36. Puesto que en la ecuación 36 están incluidas las medias de cada grupo,  $\bar{x}_i$ , se estaría considerando el efecto de la repetitividad dentro del efecto del factor. Por ello, se debe calcular una varianza neta “entre grupos” que dependa sólo de la varianza entre las medias de los grupos y la media global. La varianza neta “entre grupos se calcula como:

$$s_B^2 = \frac{s_{Factor}^2 - s_{PE}^2}{n} \quad (44)$$

La reproducibilidad se calcula finalmente como la suma de la repetitividad (error puro) y la varianza neta debida al factor (neta entre grupos):

$$s_R^2 = s_{PE}^2 + s_B^2 \quad (45)$$

Para comparar la tolerancia correspondiente a esta precisión intermedia se puede usar la conocida como Trompeta de Horwitz [11].



Horwitz dedujo una expresión para predecir el valor esperado para la desviación estándar relativa para la precisión intermedia (o interlaboratorio) a partir de la concentración de analito,  $c$  (en tanto por 1).

$$RSD_H = 2^{(1-0.5\log c)} \quad (46)$$

El valor de RSD obtenido se compara con el predicho por Horwitz ( $RSD_H$ ) mediante el parámetro Horrat [12].

$$Horrat = \frac{RSD_R}{RSD_H} \quad (47)$$

Si el valor del parámetro Horrat es igual o menor a 2 se puede decir que el método tiene valores aceptables de precisión intermedia.

Además de comparar con los valores de Horwitz, se puede también comparar con los valores establecidos por la AOAC [13] (Asociación Oficial de Químicos Analíticos). Estos valores y los obtenidos por Horwitz se presentan en la siguiente tabla.

Analyte (%)	Analyte fraction	Unit	Horwitz %RSD	AOAC PVM %RSD
100	1	100%	2	1.3
10	10 <sup>-1</sup>	10%	2.8	1.8
1	10 <sup>-2</sup>	1%	4	2.7
0.1	10 <sup>-3</sup>	0.1%	5.7	3.7
0.01	10 <sup>-4</sup>	100 ppm	8	5.3
0.001	10 <sup>-5</sup>	10 ppm	11.3	7.3
0.0001	10 <sup>-6</sup>	1 ppm	16	11
0.00001	10 <sup>-7</sup>	100 ppb	22.6	15
0.000001	10 <sup>-8</sup>	10 ppb	32	21
0.0000001	10 <sup>-9</sup>	1 ppb	45.3	30

Algunos requisitos prácticos relacionados con los estudios interlaboratorio son los siguientes [14]:

- Repetitividad:  $RSD_r = 0.5-0.6 * RSD_H$
- Reproducibilidad:  $RSD_R = 0.5-2 * RSD_H$

Para validaciones intralaboratorio [15]:

- Repetitividad:  $RSD_r = 0.2-0.3 * RSD_H$
- Reproducibilidad (precisión intermedia):  $RSD_R = 0.2-1 * RSD_H$

#### 3.4.4. Precisión y veracidad a partir del ANOVA.

A partir de un mismo diseño de ANOVA es posible estudiar tanto la precisión (veáse el apartado anterior) como la veracidad. Simplemente hay que considerar que cada medida  $x_{ij}$  tendrá un sesgo

$$\delta_{ij} = x_{ij} - T \quad (48)$$

Siendo T el valor certificado de un CRM (pues el valor verdadero  $\mu$  no se conoce).

El sesgo se puede entonces calcular como:

$$\delta = \frac{1}{hn} \sum_i^h \sum_j^n x_{ij} - T = \bar{x} - T \quad (49)$$

La incertidumbre de este sesgo se puede calcular como [16]:

$$u^2(\delta) = \frac{s_R^2 - s_r^2 \left( \frac{n-1}{n} \right)}{h} \quad (50)$$

Donde  $s_R^2$  es la varianza de reproducibilidad (puesto que consideramos un ensayo interlaboratorio), que puede ser sustituida por la varianza de precisión intermedia  $s_{IP}^2$  en caso de estar realizando la validación intralaboratorio. Se calcula a partir de la ecuación 43. La varianza de repetitividad  $s_r^2$  no es otra que la varianza de error puro  $s_{PE}^2$  obtenida en el ANOVA. Se considera que hay ausencia de sesgo para  $\frac{\delta}{u(\delta)} < 2$ .

#### 3.4.5. Ensayos de recuperación.

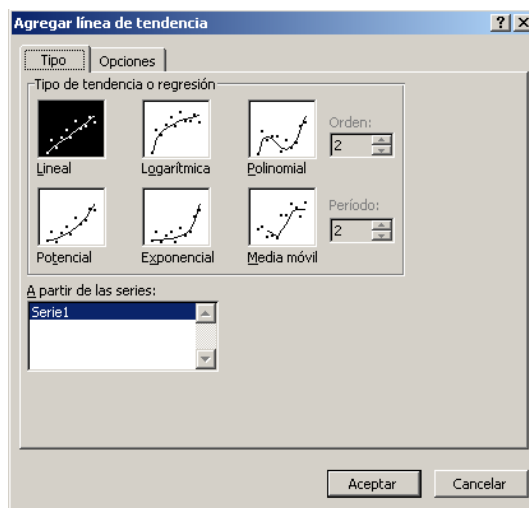
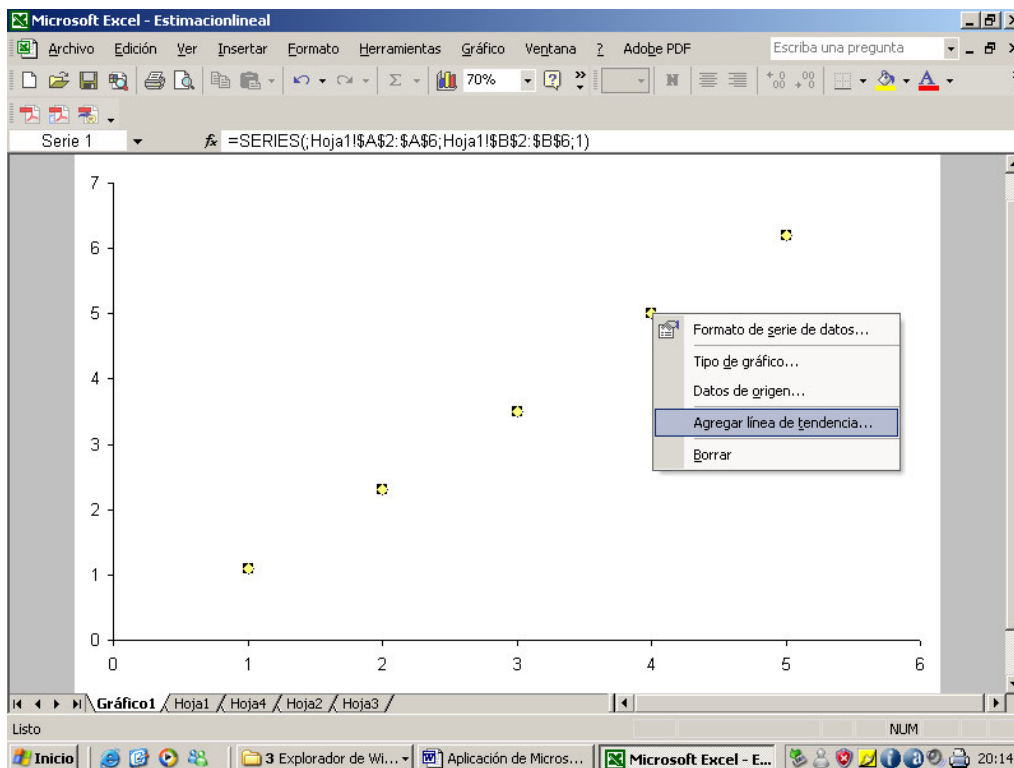
Al desarrollar un método totalmente nuevo en una matriz no muy estudiada puede ocurrir que no existan materiales de referencia ni otros métodos validados para comparar el método en desarrollo. Entonces se hace necesario el empleo de ensayos de recuperación [3]. Estos consisten en añadir una cantidad de analito conocida a una matriz que simule a la muestra, analizarla y ver la cantidad de analito que se recupera. Se deben ensayar al menos tres niveles de concentración, uno bajo, otro medio y otro alto. A veces, cuando no es posible preparar una matriz similar a la muestra se realizan los ensayos sobre la propia muestra. Se calcula la recuperación como:

$$R = \frac{C_{obs} - C_{native}}{C_{spike}} \quad (51)$$

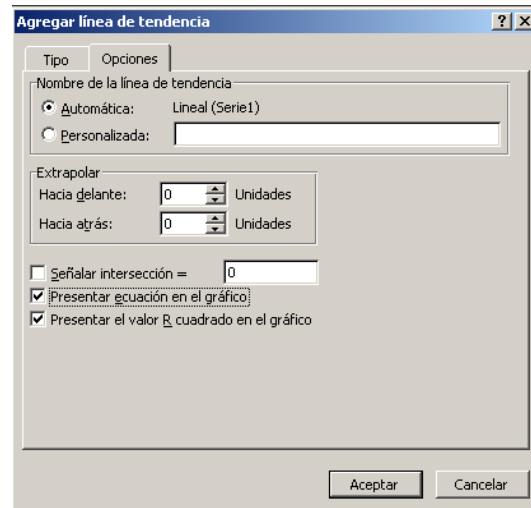
Donde  $C_{obs}$  es la concentración medida con nuestro método,  $C_{native}$  es la concentración que tiene la muestra (en el caso de ser una matriz preparada debe ser cero) y  $C_{spike}$  es la cantidad que hemos añadido. Si realizamos  $j$  replicados en cada uno de los  $i$  niveles podemos ver si las recuperaciones son equivalentes en cada uno de los niveles estudiados.

### 3.5. Regresión

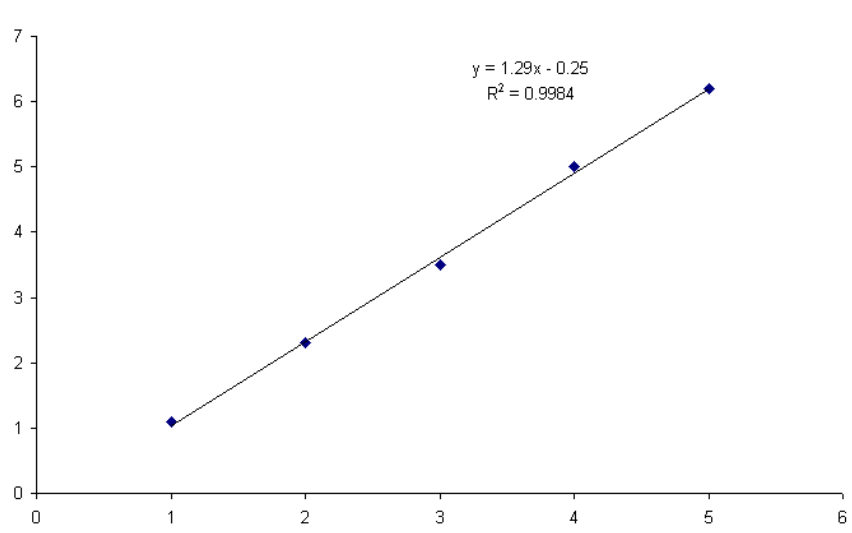
Ya se ha explicado la forma de calcular los coeficientes de regresión, cuando la relación entre las variables  $x$  e  $y$  es lineal. Por todos es conocido que si representamos los datos en Excel y seleccionamos los puntos, pulsando en el botón secundario del ratón, en “Agregar línea de tendencia”, como se indica en la figura, se puede obtener la recta de regresión con su ecuación correspondiente. Los pasos a seguir son los siguientes:



Se selecciona el tipo de estimación y se pulsa en la pestaña “Opciones”. Marcar “Presentar ecuación en el gráfico” y “Presentar valor de R cuadrado en el gráfico”. Pulsar en aceptar.



El resultado es el siguiente:



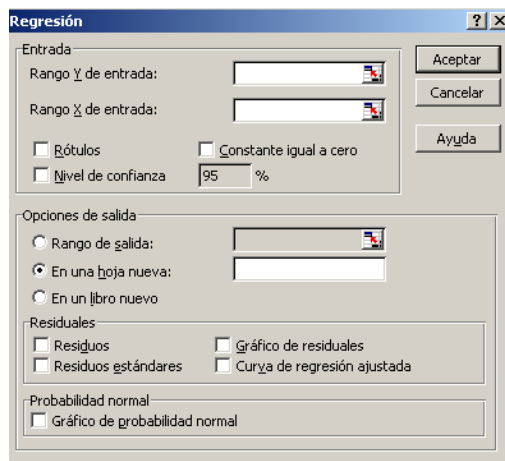
Ahora vamos a explicar la herramienta “Regresión” del menú Herramientas>Análisis de datos.

Partimos de los datos:

	A	B
1	x	y
2	1	1.1
3	2	2.3
4	3	3.5
5	4	5
6	5	6.2

El formulario de la herramienta Regresión es el siguiente:





Los casillas de rangos de entrada y salida y la de rótulos ya son conocidas por nosotros. La casilla “Constante igual a cero” se utiliza si queremos forzar a la curva a pasar por el cero. Nosotros no la forzaremos. Esta herramienta ofrece siempre el intervalo de confianza de la pendiente y la ordenada en el origen al 95% de nivel de confianza, si se desea otro nivel de confianza se puede activar la casilla “Nivel de confianza, por ejemplo para el 90%. En el cuadro “Residuales” se puede especificar si se quieren tener los residuales, residuales estandarizados, el gráfico de residuales y el de la propia curva ajustada. Los residuos estándares se obtienen como:

$$Res_{st,i} = \frac{Res_i - Res_{mean}}{Res_{sd}} \quad (52)$$

Donde  $Res_{st,i}$  es el  $i$ -ésimo residual estandarizado,  $Res_i$  el  $i$ -ésimo residual,  $Res_{mean}$  es la media de los  $i$  residuales y  $Res_{sd}$  su desviación estándar.

Finalmente se puede obtener la curva de probabilidad normal. Este último presenta los valores de  $y$  ordenados de menor a mayor indicando el percentil al que pertenece. Si los puntos parecen ajustarse a una línea recta, puede decirse que parece indicar que los datos provienen de una distribución normal, pero téngase en cuenta que *en algunos casos, aunque los puntos se ajusten a una línea recta puede que los datos no provengan de una distribución normal*; por ello se recomienda utilizar métodos objetivos.

Seleccionamos B1:B9 para el rango  $y$  y A1:A9 para el  $x$ . Seleccionamos Rótulos, Residuos, Gráfico de residuales y Curva de regresión ajustada. El resultado será el siguiente:

Estadísticas de la regresión	
Coefficiente de correlación múltiple	0.999189737
Coefficiente de determinación R <sup>2</sup>	0.99838013
R <sup>2</sup> ajustado	0.997840173
Error típico	0.09486833
Observaciones	5

En las estadísticas de regresión se tiene el coeficiente de correlación y el error típico ( $s_{y/x}$ ). El coeficiente de determinación  $R^2$ , que coincide con el cuadrado del coeficiente de correlación, se obtiene a partir de la suma de cuadrados de los residuales (SCRes) y la suma de cuadrados total (SCT):

$$SCRes = \sum_i^n (y_i - \hat{y}_i)^2 \quad (53)$$

$$SCT = \sum_i^n (y_i - \bar{y})^2 \quad (54)$$

$$R^2 = 1 - \frac{SCRes}{SCT} \quad (55)$$

El ajustado se calcula a partir de las varianzas de los residuales y total, también llamados cuadrados medios o promedio, CMRes y CMT.

$$CMRes = \frac{\sum_i^n (y_i - \hat{y}_i)^2}{n-2} \quad (56)$$

$$CMT = \frac{\sum_i^n (y_i - \bar{y})^2}{n-1} \quad (57)$$

$$R^{2'} = 1 - \frac{CMRes}{CMT} \quad (58)$$

Cuanto más próximo a 1 sea el coeficiente de correlación mejor será el ajuste lineal. No es el criterio que tomaremos, trataremos este tema cuando hablemos de linealidad y rango lineal.

ANÁLISIS DE VARIANZA					
	Grados de libertad	Suma de cuadrados	Promedio de los cuadrados	F	Valor crítico de F
Regresión	1	16.641	16.641	1849	2.76835E-05
Residuos	3	0.027	0.009		
Total	4	16.668			

El análisis de la varianza aquí realizado se basa en que la suma de cuadrados totales se descompone en dos componentes, la suma de cuadrados de los residuales y la de regresión:

$$\sum_i^n (y_i - \bar{y})^2 = \sum_i^n (\hat{y}_i - \bar{y})^2 + \sum_i^n (y_i - \hat{y}_i)^2 \quad (59)$$

Estas sumas de cuadrados aparecen en la tercera columna y dividiendo entre los grados de libertad de la segunda se obtienen las correspondientes varianzas (llamadas por EXCEL promedio de cuadrados). Excel solo calcula las varianzas de regresión y la de los residuales (cuarta columna). A partir de ellas calcula un Valor de F (quinta columna) y calcula una probabilidad para este valor (es lo que llama valor crítico de F).

$$F = \frac{\frac{\sum_i^n (\hat{y}_i - \bar{y})^2}{(n-1) - (n-2)}}{\frac{\sum_i^n (y_i - \hat{y}_i)^2}{n-2}} \quad (60)$$

En general, si el valor de F es muy elevado, la suma de cuadrados de residuales será mucho menor que la de regresión, con lo que también será mucho menor que la suma de cuadrados totales. Entonces, según la expresión 55,  $r$  será próximo a 1.

El hecho de que EXCEL trabaje con un valor de probabilidad es similar a comparar con el valor tabulado de F para 1 y  $n-2$  grados de libertad, pero en vez de comparar, calcula directamente la probabilidad de que el F calculado sea mayor que el tabulado. Si la probabilidad es muy pequeña es poco probable que los datos lleven a un valor tan elevado de F, es decir, es poco probable que la suma de cuadrados totales sea mucho mayor que la de residuales y  $r$  sea próximo a uno. Esto quiere decir que el ajuste lineal no es fruto de la casualidad, porque es poco probable que los datos se organicen de este modo por casualidad.

Como ya hemos dicho, nosotros estamos más acostumbrados a comparar el valor de F experimental con un valor tabulado de F para un nivel de confianza determinado (95 %, por ejemplo) teniendo en cuenta que los grados de libertad para la varianza de regresión es  $(n-1-n-2=1)$  y la de la varianza de los residuales  $n-2$ . Si  $F_{\text{exp}} > F(0.05, 1, n-2)$  la correlación lineal es significativa y no puede atribuirse a la casualidad.

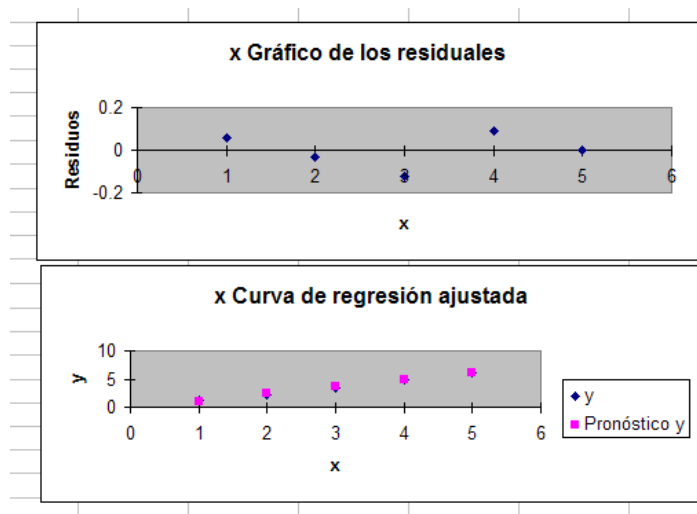
Los parámetros ajustados de pendiente y ordenada se obtienen en:

16						
17	Intercepción	-0.25	0.099498744	-2.512594538	0.086739396	-0.566649706
18	x	1.29	0.03	43	2.76835E-05	1.194526521
19						1.385473479

Intercepción se refiere a la ordenada y X a la pendiente. En la columna de error típico aparecen las desviaciones estándar de la ordenada y la pendiente. Se calcula un valor de  $t$  como el valor del parámetro dividido de su desviación estándar, comparando con  $t(0.05, n-2)$  grados de libertad. Igual que en el caso anterior se tiene la probabilidad de que éste  $t$  experimental sea mayor que el teórico. Los límites inferiores y superiores se calculan restando y sumando al valor del parámetro el valor obtenido al multiplicar su error por  $t(0.05, n-2)$ .

También se observan los valores de los residuales y los gráficos de residuales y de la curva ajustada.

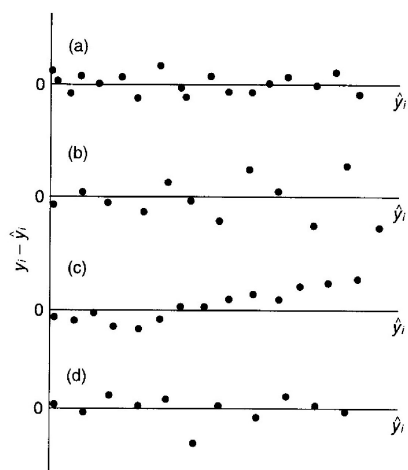
22	Análisis de los residuales		
23			
24	Observación	Pronóstico y	Residuos
25	1	1.04	0.06
26	2	2.33	-0.03
27	3	3.62	-0.12
28	4	4.91	0.09
29	5	6.2	8.88178E-16
30			



Observando los residuales podemos ver si:

- Si el modelo se ajusta bien a los datos. Los residuos se distribuyen homogéneamente.
- Si los datos son heterocedásticos, es decir, que la varianza no es constante en todo el rango de  $x$ .
- Si el modelo es inadecuado. Se observa cuando los residuales de valores próximos en  $x$  son todos del mismo signo, cambiando de signo en otro rango de  $x$ .

d) Si existen outliers.



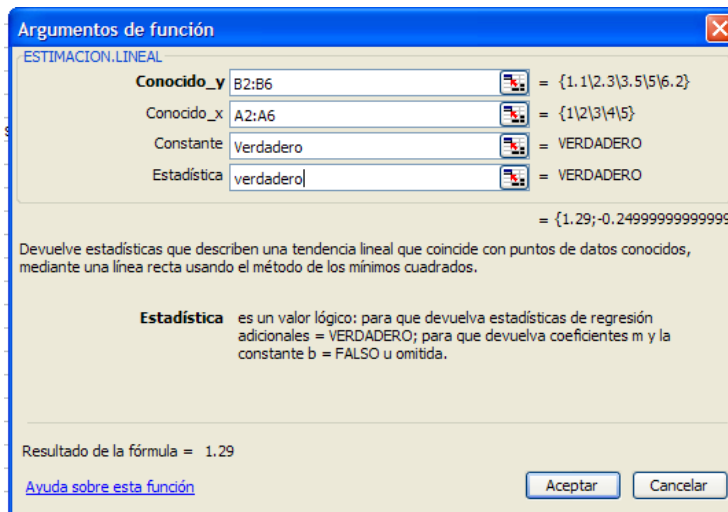
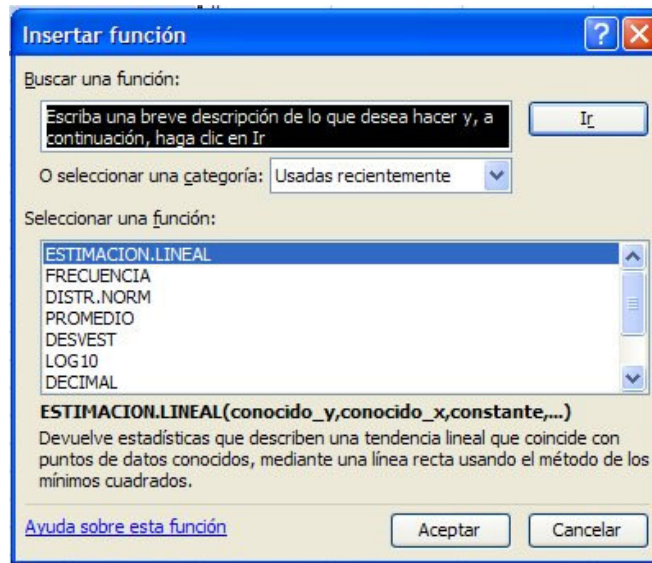
### 3.5.1. La función =ESTIMACION.LINEAL()

Esta función se utiliza de una forma especial, puesto que se trata de una formula matricial. En primer lugar hay que seleccionar un rango de 10 celdas tal y como se indica en la figura. En esta figura se ha escrito en las celdas adyacentes a las del rango el parámetro que la formula devolverá.

	A	B	C	D	E	F
1	x	y				
2	1	1.1	b			a
3	2	2.3	sb			sa
4	3	3.5	r2			sy/x
5	4	5	F			df
6	5	6.2	SCReg			SCRes

La pendiente  $b$  y su desviación estándar  $sb$ , la ordenada  $a$  y su desviación estándar  $sa$ , el coeficiente de correlación al cuadrado  $r^2$  (para obtener  $r$  realizar la raíz cuadrada) y la desviación estándar de regresión  $sy/x$  son los parámetros fundamentales que se han explicado en la sección 1.7. El valor de  $F$  se corresponde con el anteriormente descrito y  $df$  son los grados de libertad para los residuales. El parámetro aquí llamado  $ssres$  es la suma de los residuales al cuadrado (Ecuación 7). En cuanto a  $ssreg$ , es la suma de cuadrados de regresión.

El siguiente paso es seleccionar la fórmula ESTIMACION.LINEAL del menú “Insertar /Función” de EXCEL



En “Conocido\_y” se debe seleccionar los valores de y, en “Conocido\_x” los de x. En “Constante” se debe poner un 1 o VERDADERO para que calcule la ordenada en el origen. En “Estadística” se pondrá un 1 o VERDADERO para que el calculo de ssreg y ssres sea de la forma anteriormente explicada.

En vez de pulsar aceptar, sino CTRL+MAYÚS+ENTRAR, para que la fórmula de resultados en todas las celdas de la matriz.

	A	B	C	D	E	F
1	x	y				
2		1	1.1	b	1.29	-0.25
3		2	2.3	sb	0.03	0.09949874
4		3	3.5	r <sup>2</sup>	0.99838013	0.09486833
5		4	5	F	1849	3
6		5	6.2	SCReg	16.641	0.027
7						

### 3.5.2. Linealidad: Prueba de la falta de ajuste (lack of fit).

Se realiza esta prueba para ver si el modelo se ajusta bien a los datos. Como requisito indispensable haber realizado replicados por cada nivel de concentración. En líneas generales se trata de corregir la varianza de los residuales con la varianza de repetición pura promediada para todos los niveles y comprobar si esta varianza corregida (la varianza de falta de ajuste  $s_{LOF}^2$ ) difiere de la de repetición pura mediante un test F.

Para poder promediar las varianzas de repetición en cada nivel de concentración es necesario que estas no difieran entre sí. Para ello se realizaría el test de Cochran para detección de outliers dispersivos.

Supongamos que realizamos  $r$  replicados en  $n$  niveles de concentración. Cuando se tiene una serie de resultados agrupados en replicados, y se sospecha que una serie de replicados puede constituir un outlier dispersivo, es decir, que presenta una varianza sospechosa, por ser la mayor de la serie, se puede usar la prueba de Cochran [17]. De esta forma se decide si dicha varianza puede considerarse o no un outlier. El estadístico  $C$  de Cochran se calcula según la siguiente fórmula:

$$C = \frac{s_{\max}^2}{\sum_{i=1}^h s_i^2} \quad (61)$$

El valor obtenido se compara con el tabulado  $C(h, r, \alpha)$ . Siendo  $s_{\max}^2$  la varianza sospechosa,  $h$  el número de series de replicados,  $s_i^2$  las varianzas de cada una de estas series,  $r$  el número de replicados y  $\alpha$  el nivel de significación. Si el valor calculado es mayor que el teórico, normalmente para un 95% de nivel de confianza, se considera que la serie de replicados sospechosa constituye un outlier dispersivo. En este caso no tiene sentido la prueba de la falta de ajuste, puesto que los datos no son homocedásticos y no se debería aplicar la regresión lineal ordinaria, sino la ponderada. En el ANEXO 2 se presentan los valores tabulados del parámetro  $C$ .

Una vez hecho esto, si no hay outliers dispersivos, se calcula una varianza de error puro como la media de las varianzas a cada nivel (suponiendo  $h$  niveles).

$$s_{PE}^2 = \frac{\sum_{i=1}^h s_i^2}{h} \quad (62)$$

Con  $r-1$  grados de libertad (se han sumado  $h$  varianzas de  $r-1$  grados de libertad pero se ha dividido entre  $h$ ).

A partir del ANOVA se tiene la varianza de los residuales:

$$s_{RES}^2 = \frac{\sum_i^n (\hat{y}_i - y_i)^2}{n-2} \quad (63)$$

Para obtener la varianza de falta de ajuste se resta la de error puro a la de los residuales.

$$s_{LOF}^2 = \frac{(n-2)s_{RES}^2 - (r-1)s_{PE}^2}{(n-2)-(r-1)} \quad (64)$$

Tras esto se calcula el factor  $F$  de falta de ajuste:

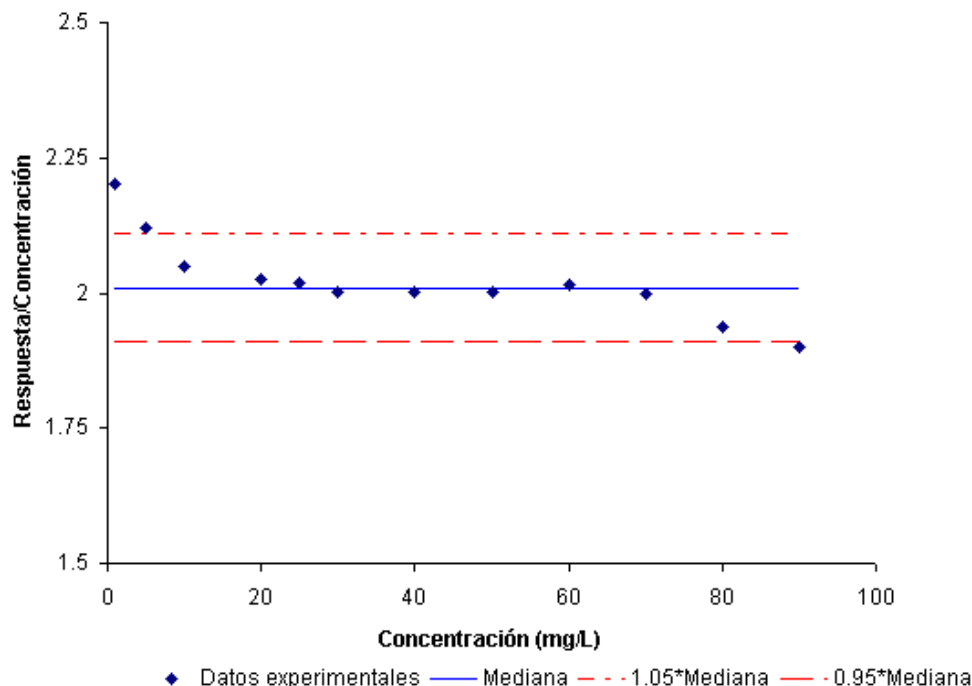
$$F_{LOF} = \frac{s_{LOF}^2}{s_{PE}^2} \quad (65)$$

Si  $F_{LOF} < F_{crit}(0.05, n-r-1, r-1)$  la varianza de falta de ajuste no será estadísticamente mayor que la de error puro. Esto quiere decir que no existe un error debido a la falta de ajuste del modelo que no pueda ser atribuido al propiamente aleatorio, por lo tanto, el modelo lineal se ajusta perfectamente a los datos experimentales.

### 3.5.3. Linealidad e intervalo dinámico: método del factor de respuesta de Huber.

El método del factor de respuesta de Huber [18] se aplica cuando tenemos rectas de calibrado corregidas con el blanco para que la ordenada sea prácticamente cero. Así, si la respuesta es prácticamente  $y=bx$ , los factores de respuesta  $R=y/x$  serán próximos a  $b$ . Se representa  $y/x$  frente a la concentración de analito (o frente al logaritmo si el intervalo de trabajo es amplio, por ejemplo 1-1000 mg/L). Calculamos la mediana de los factores de respuesta,  $M$ , trazando una recta en este valor y dos rectas más con valores 1.05M y 0.95M. Los puntos experimentales que se encuentren entre esas dos líneas corresponden al intervalo lineal.





#### 3.5.4. Comparación con un método de referencia.

Para comparar un método que está siendo validado con uno de referencia usando una recta de regresión se representan los resultados obtenidos por el método a validar frente a los obtenidos por el de referencia, obteniendo la ecuación de la recta  $y = (b \pm s_b)x + (a \pm s_a)$ . Donde  $y$  son los resultados obtenidos con el método a validar y  $x$  los obtenidos por el de referencia. Si los dos métodos llevan a resultados idénticos, es obvio que la pendiente debe ser igual a la unidad y la ordenada igual a cero. Se calculan los parámetros  $t$  de Student:

$$t_a = \frac{|a-0|}{s_a} \quad (66)$$

$$t_b = \frac{|b-1|}{s_b} \quad (67)$$

Se comparan con el valor tabulado  $t(0.05, n-2)$ .

Si  $b=1$  y  $a=0$ , ambos métodos llevan a resultados estadísticamente iguales. Los métodos son equivalentes.

Si  $b=1$  y  $a \neq 0$ , el método a calibrar da un resultado más alto o más bajo que el de referencia, pero en una cantidad fija. es decir, se producen errores sistemáticos constantes o aditivos.

Si  $b \neq 1$  y  $a = 0$ , se produce un error sistemático proporcional a la concentración de analito en el método a validar.

Si  $b \neq 0$  y  $a \neq 0$ , existen los dos tipos de errores.

### 3.5.6. Efecto matriz.

El efecto matriz se debe, como su propio nombre indica a la matriz de la muestra. Las muestras, salvo en contadas ocasiones, deben tratarse antes de ser medidas. La medida se realiza con una técnica determinada y la señal obtenida se interpola en los valores obtenidos mediante una recta de calibrado. A veces, el hecho de tratar la muestra no es suficiente, pues puede quedar en una matriz que difiere de los estándares utilizados para realizar el calibrado y es muy posible que esa matriz interfiera en el análisis. Se debe estudiar entonces si existe o no dicho efecto. Si es así, no podremos realizar el el análisis utilizando un calibrado externo, sino que deberemos usar el método de adición patrón.

Se puede detectar la presencia de efecto matriz comparando el calibrado externo con el de adición patrón. Se trataría de obtener la pendiente de calibrado externo y la de adición patrón y compararlas. En primer lugar realizamos el cociente entre ambas pendientes y calculamos la desviación estándar asociada al mismo:

$$R = \frac{b_{CE}}{b_{AP}} \quad (68)$$

$$S_R = R \sqrt{\left( \frac{s_{b_{CE}}^2}{b_{CE}^2} + \frac{s_{b_{AP}}^2}{b_{AP}^2} \right)} \quad (69)$$

Donde  $b_{CE}$  y  $b_{AP}$  son las pendientes de las rectas de calibrado externo y adición patrón, respectivamente,  $s_{b_{CE}}$  y  $s_{b_{AP}}$  son sus desviaciones estándar correspondientes. Si ambas pendientes son prácticamente iguales, su cociente,  $R$ , no debe ser significativamente distinto de 1. Calculando la  $t$  de Student como:

$$t_{cal} = \frac{|R-1|}{S_R} \quad (70)$$

El valor calculado se compara con el factor de cobertura  $k = 2$  para un 95 % de nivel de confianza [10]. Si  $t_{cal}$  es mayor que 2, las dos pendientes no serán comparables y el análisis deberá realizarse por el método de adición patrón.

**Bibliografía**

1. The fitness for purpose of analytical methods. A laboratory guide to method validation and related topics. EURACHEM, LGC, Teddington, UK 1998. <http://www.eurachem.org/guides/valid.pdf>.
2. R. Compañó, A. Rios. Garantía de la calidad en los laboratorios analíticos. Editorial Síntesis, Madrid, 2002.
3. J. M. Jurado, A. Alcázar, F. Pablos, M. J. Martín, A. G. González. *Talanta*, **2004**, 63, 297-302
4. J. Inczedy, T. Lengyel, A. M. Ure. Compendium of Analytical Nomenclature 3rd edition. Blackwell Science, 1998
5. L. Cuadros, A.M. García and J.M. Bosque, *Anal. Lett.* **1996**, 29, 1231
6. L. Huber. *LC-GC Int.* **1998**, 11, 96–105
7. J. N. Miller, J. C. Miller. Estadística y quimiometría para química analítica. Prentice Hall, Madrid, 2002.
8. J. A. Rafter, M. L. Abell, J. P. Braselton. Multiple comparison methods for means. *SIAM Review*, **2002**, 44, 259-278.
9. H. L. Harter. Tables of range and studentized range. *Annals of Mathematical Statistics*, **1960**, 31, 1122-1147.
10. A. G. González, M. A. Herrador. A practical guide to analytical method validation, including measurement uncertainty and accuracy profiles. *Trends in Analytical Chemistry*, **2007**, 26, 227-238.
11. W. Horwitz. Evaluation of Analytical Methods used for regulations of foods and drugs. *Analytical Chemistry*, **1982**, 54, 67A-76A.
12. R. Wood. How to validate analytical methods. *Trends in Analytical Chemistry*, **1999**, 18, 624-632.
13. AOAC International. Method Validation Program (OMA/PVM Department), including Appendix D: Guidance for collaborative study procedures to validate characteristics of a method of analysis, **2000**

14. I. Taverniers, M. De Loose, E. van Bockstaele, *Trends in Analytical Chemistry*, **2004**, 23, 535.
15. E. Pitchard, *Quality in the analytical chemistry laboratory*. ACOL Series, Wiley, Chichester, West Sussex, UK, **1995**.
16. International Organization of Standardization (ISO), ISO/DTS 21748, *Guide to the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation*, ISO, Geneva, Switzerland, **2003**
17. W. G. Cochran. The distribution of the largest of a set of estimated variances as a fraction of their total. *Annals of Eugenics*, **1941**, 11, 47-61.
18. L. Huber. Validation of analytical methods: review and strategy. *LC-GC International*. **1998**, 11, 96-105.

ANEXO 1. Tabla de valores estandarizados para el test de Tukey

Studentized Range 0.95 Quantiles,  $q_{0.95}(r, \nu)$

Number of variates (r)													
$\nu$	2	3	4	5	6	7	8	9	10	11	12	13	$\nu$
1	17.97	26.98	32.82	37.08	40.41	43.12	45.40	47.36	49.07	50.59	51.96	53.20	1
2	6.085	8.331	9.799	10.88	11.73	12.43	13.03	13.54	13.99	14.40	14.76	15.09	2
3	4.501	5.910	6.825	7.502	8.037	8.478	8.852	9.177	9.462	9.717	9.946	10.15	3
4	3.926	5.040	5.757	6.287	6.706	7.053	7.347	7.602	7.826	8.027	8.208	8.373	4
5	3.635	4.602	5.218	5.673	6.033	6.330	6.582	6.801	6.995	7.167	7.323	7.466	5
6	3.460	4.339	4.896	5.305	5.628	5.895	6.122	6.319	6.493	6.649	6.789	6.917	6
7	3.344	4.165	4.681	5.060	5.359	5.606	5.815	5.997	6.158	6.302	6.431	6.550	7
8	3.261	4.041	4.529	4.886	5.167	5.399	5.596	5.767	5.918	6.053	6.175	6.287	8
9	3.199	3.948	4.415	4.755	5.024	5.244	5.432	5.595	5.738	5.867	5.983	6.089	9
10	3.151	3.877	4.327	4.654	4.912	5.124	5.304	5.460	5.598	5.722	5.833	5.935	10
11	3.113	3.820	4.256	4.574	4.823	5.028	5.202	5.353	5.486	5.605	5.713	5.811	11
12	3.081	3.773	4.199	4.508	4.750	4.950	5.119	5.265	5.395	5.510	5.615	5.710	12
13	3.055	3.734	4.151	4.453	4.690	4.884	5.049	5.192	5.318	5.431	5.533	5.625	13
14	3.033	3.701	4.111	4.407	4.639	4.829	4.990	5.130	5.253	5.364	5.463	5.554	14
15	3.014	3.673	4.076	4.367	4.595	4.782	4.940	5.077	5.198	5.306	5.403	5.492	15
16	2.998	3.649	4.046	4.333	4.557	4.741	4.896	5.031	5.150	5.256	5.352	5.439	16
17	2.984	3.628	4.020	4.303	4.524	4.705	4.858	4.991	5.108	5.212	5.306	5.392	17
18	2.971	3.609	3.997	4.276	4.494	4.673	4.824	4.955	5.071	5.173	5.266	5.351	18
19	2.960	3.593	3.977	4.253	4.468	4.645	4.794	4.924	5.037	5.139	5.231	5.314	19
20	2.950	3.578	3.958	4.232	4.445	4.620	4.768	4.895	5.008	5.108	5.199	5.282	20
24	2.919	3.532	3.901	4.166	4.373	4.541	4.684	4.807	4.915	5.012	5.099	5.179	24
30	2.888	3.485	3.845	4.102	4.301	4.464	4.601	4.720	4.824	4.917	5.001	5.077	30
40	2.858	3.442	3.791	4.039	4.232	4.388	4.521	4.634	4.735	4.824	4.904	4.977	40
60	2.829	3.399	3.737	3.977	4.163	4.314	4.441	4.550	4.646	4.732	4.808	4.878	60
120	2.800	3.356	3.685	3.917	4.096	4.241	4.363	4.468	4.560	4.641	4.714	4.781	120
$\infty$	2.772	3.314	3.633	3.858	4.030	4.170	4.286	4.387	4.474	4.552	4.622	4.685	$\infty$

Number of variates (r)													
$\nu$	14	15	16	17	18	19	20	30	40	60	80	100	$\nu$
1	54.33	55.36	56.32	57.22	58.04	58.83	59.56	65.15	68.92	73.97	77.40	79.98	1
2	15.39	15.65	15.92	16.14	16.38	16.57	16.78	18.27	19.28	20.66	21.59	22.29	2
3	10.35	10.52	10.69	10.84	10.98	11.11	11.24	12.21	12.86	13.76	14.36	14.82	3
4	8.524	8.664	8.793	8.914	9.027	9.133	9.233	10.00	10.53	11.24	11.73	12.10	4
5	7.596	7.716	7.828	7.932	8.030	8.122	8.208	8.875	9.330	9.949	10.37	10.69	5
6	7.034	7.143	7.244	7.338	7.426	7.508	7.586	8.189	8.601	9.162	9.547	9.839	6
7	6.658	6.759	6.852	6.939	7.020	7.097	7.169	7.727	8.110	8.631	8.989	9.260	7
8	6.389	6.483	6.571	6.653	6.729	6.801	6.869	7.395	7.756	8.248	8.586	8.843	8
9	6.186	6.275	6.359	6.437	6.510	6.579	6.643	7.144	7.488	7.958	8.281	8.526	9
10	6.028	6.114	6.194	6.269	6.339	6.405	6.467	6.948	7.278	7.730	8.041	8.276	10
11	5.901	5.984	6.062	6.134	6.202	6.265	6.325	6.790	7.109	7.546	7.847	8.075	11
12	5.797	5.878	5.953	6.023	6.089	6.151	6.209	6.660	6.970	7.394	7.687	7.908	12
13	5.711	5.789	5.862	5.931	5.995	6.055	6.112	6.551	6.853	7.267	7.552	7.769	13
14	5.637	5.714	5.785	5.852	5.915	5.973	6.029	6.459	6.754	7.159	7.437	7.649	14
15	5.574	5.649	5.719	5.785	5.846	5.904	5.958	6.379	6.669	7.065	7.338	7.546	15
16	5.519	5.593	5.662	5.726	5.786	5.843	5.896	6.310	6.594	6.983	7.252	7.456	16
17	5.471	5.544	5.612	5.675	5.734	5.790	5.842	6.249	6.529	6.912	7.176	7.377	17
18	5.429	5.501	5.567	5.629	5.688	5.743	5.794	6.195	6.471	6.848	7.108	7.307	18
19	5.391	5.462	5.528	5.589	5.647	5.701	5.752	6.147	6.419	6.791	7.048	7.244	19
20	5.357	5.427	5.492	5.553	5.610	5.663	5.714	6.104	6.372	6.740	6.994	7.187	20
24	5.251	5.319	5.381	5.439	5.494	5.545	5.594	5.968	6.226	6.578	6.822	7.007	24
30	5.147	5.211	5.271	5.327	5.379	5.429	5.475	5.833	6.080	6.417	6.650	6.827	30
40	5.044	5.106	5.163	5.216	5.266	5.313	5.358	5.700	5.934	6.255	6.477	6.645	40
60	4.942	5.001	5.056	5.107	5.154	5.199	5.241	5.566	5.789	6.093	6.302	6.462	60
120	4.842	4.898	4.950	4.998	5.043	5.086	5.126	5.434	5.644	5.929	6.126	6.275	120
$\infty$	4.743	4.796	4.845	4.891	4.934	4.974	5.012	5.301	5.498	5.764	5.947	6.085	$\infty$

ANEXO 2. Parámetro C para el contraste de Cochran. k: conjunto de datos, n:  
replicados. 95% n.c.

k	n						
	2	3	4	5	6	7	8
3	0,9985	0,9750	0,9392	0,9057	0,8772	0,8534	0,8332
3	0,9669	0,8709	0,7977	0,7457	0,7071	0,6771	0,6530
4	0,9066	0,7679	0,6841	0,6287	0,5896	0,5598	0,5365
5	0,8412	0,6838	0,5981	0,5441	0,5066	0,4783	0,4564
6	0,7808	0,6161	0,5321	0,4803	0,4447	0,4184	0,3980
7	0,7271	0,5612	0,4800	0,4307	0,3974	0,3728	0,3535
8	0,6798	0,5157	0,4377	0,3910	0,3596	0,3362	0,3188
9	0,6385	0,4775	0,4027	0,3584	0,3296	0,3067	0,2901
10	0,6020	0,4450	0,3733	0,3311	0,3029	0,2823	0,2666
12	0,5410	0,3924	0,3284	0,2890	0,2624	0,2439	0,2299
15	0,4709	0,3348	0,2758	0,2419	0,2195	0,2034	0,1911
20	0,3884	0,2705	0,2205	0,1921	0,1735	0,1602	0,1501
24	0,3434	0,2354	0,1907	0,1656	0,1493	0,1374	0,1286
30	0,2929	0,1980	0,1598	0,1377	0,1237	0,1137	0,1061
40	0,2370	0,1576	0,1259	0,1082	0,0968	0,0887	0,0827
60	0,1737	0,1131	0,0895	0,0765	0,0682	0,0623	0,0563
120	0,0996	0,0682	0,0495	0,0419	0,0371	0,0337	0,0312
∞	0	0	0	0	0	0	0

observaciones de cada conjunto

9	10	11	17	37	145	∞
0,8159	0,8010	0,7890	0,7341	0,6602	0,5813	0,5000
0,8333	0,8167	0,8025	0,5465	0,4748	0,4031	0,3333
0,5175	0,5017	0,4884	0,4366	0,3720	0,3093	0,2500
9,4387	0,4241	0,4118	0,3645	0,3066	0,2513	0,2000
0,3817	0,3682	0,3568	0,3135	0,2612	0,2119	0,1667
0,3384	0,3259	0,3154	0,2756	0,2278	0,1833	0,1429
0,3043	0,2926	0,2829	0,2462	0,2022	0,1616	0,1250
0,2768	0,2659	0,2568	0,2226	0,1820	0,1446	0,1111
0,2541	0,2439	0,2353	0,2032	0,1655	0,1308	0,1000
0,2187	0,2098	0,2020	0,1737	0,1403	0,1100	0,0833
0,1815	0,1736	0,1671	0,1429	0,1144	0,0889	0,0667
0,1422	0,1357	0,1309	0,1108	0,0879	0,0675	0,0500
0,1216	0,1160	0,1113	0,0942	0,0743	0,0567	0,0417
0,1002	0,0958	0,0921	0,0771	0,0604	0,0457	0,0333
0,0780	0,0745	0,0713	0,0595	0,0462	0,0347	0,0250
0,0552	0,0520	0,0497	0,0411	0,0316	0,0234	0,0167
0,0292	0,0279	0,0266	0,0218	0,0165	0,0120	0,0083
0	0	0	0	0	0	0