MINISTERIO DE TRABAJO Y ASUNTOS SOCIALES
INSTITUTO NACIONAL DE SEGURIDAD E HIGIENE EN EL TRABAJO
Documentación

NTP 347: Contaminantes químicos: evaluación de la concentración ambiental

Contamination chimique: évaluation de la concentration dans l'air
Chemical contamination: air concentration assesment

Redactor:

Emilio Castejón Vilella
Ingeniero Industrial Licenciado en Farmacia

CENTRO NACIONAL DE CONDICIONES DE TRABAJO

Introducción

Aunque en los años cincuenta y sesenta algunos estudiosos se interesaron en el desarrollo de modelos que explicaran la variabilidad observada en los resultados de las tomas de muestras ambientales efectuadas por los higienistas industriales, no fue hasta la década de los setenta cuando el tema empezó a preocupar a los higienistas profesionales. La razón de ello fue que en 1970 el Congreso de los Estados Unidos de América aprobó la ley de Seguridad e Higiene en el Trabajo (Occupational Safety and Health Act) introduciendo la posibilidad de que las mediciones ambientales efectuadas por la Administración pudieran ser la base, si se demostraba que se superaban los límites prescritos reglamentariamente, para la imposición de sanciones a las empresas.

El problema se planteaba en los siguientes términos: puesto que la ley (americana) exige que la concentración legal de referencia sea respetada **dentro de cualquier jornada laboral**, para que un inspector pueda sancionar a la empresa es necesario que pueda **demostrar** que la concentración promedio durante la jornada en la cual se ha realizado la inspección ha superado lo legalmente admitido. Si el inspector ha efectuado una medición durante la totalidad de la jornada, el resultado obtenido en la medición sólo viene afectado por la incertidumbre inherente a los sistemas de muestreo y análisis empleados, incertidumbre que está habitualmente bien cuantificada, por lo que es sencillo calcular el intervalo dentro del cual puede encontrarse la concentración media real y concluir si existe o no infracción; pero cuando, cualesquiera que sean las razones, la toma de muestra no puede cubrir la totalidad de la jornada y, por tanto, los resultados obtenidos son representativos únicamente de una parte de aquella, aparece una nueva incógnita: ¿son los datos obtenidos durante una parte de la jornada, extrapolables a la totalidad de la misma?

Las condiciones bajo las cuales los resultados de una medición que no fuera de jornada completa podían permitir efectuar afirmaciones con respecto a la concentración media diaria y el riesgo de error de dichas afirmaciones, fueron objeto de múltiples publicaciones durante los primeros años setenta, si bien posteriormente el interés por el tema decayó, debido a que el problema era sobre todo norteamericano y originado por una legislación extremadamente peculiar.

En los últimos años, sin embargo, han aparecido una serie de nuevos estudios que replantean el problema desde una perspectiva más amplia e incluyen algunos aspectos que pueden resultar de interés general en la práctica diaria de la Higiene Industrial.

Variabilidad de las concentraciones ambientales

La elevada variabilidad de los resultados de varias tomas de muestra efectuadas en el mismo puesto de trabajo en distintos periodos de. tiempo, es bien conocida de los higienistas industriales. Dicha variabilidad es atribuible por un lado, al método experimental de toma de muestras y análisis empleado que, como es lógico, tiene errores inherentes a su propia naturaleza experimental. Este factor de variabilidad es, sin embargo, poco relevante (1) en comparación con el resto de causas que contribuyen al fenómeno y que suelen agruparse bajo el nombre de «causas ambientales"; entre ellas se encuentran las variables propias del proceso de producción (continuo, discontinuo), la variabilidad en la secuencia de operaciones que realiza el trabajador, las características cambiantes de las materias primas, los sistemas de ventilación, etc. La magnitud de la variabilidad observada es relevante tanto para los resultados de los muestreos efectuados en distintos periodos de tiempo a un mismo trabajador, como respecto a las diferencias observadas entre distintos trabajadores.

Con respecto a las diferencias observadas entre distintos trabajadores, y a fin de reducir el número de muestras necesarias para la evaluación de un colectivo, se ha propuesto la clasificación de los mismos en los llamados "grupos homogéneos de exposición" en base a las similitudes existentes entre los puestos de trabajo respectivamente ocupados. La experiencia demuestra (2) que dicho método es ineficaz y conduce a la definición de grupos homogéneos de exposición con un alto grado de heterogeneidad, debido a que las variables usualmente empleadas para definir la homogeneidad de exposición del grupo (puesto de trabajo, tipo de contaminante, naturaleza del proceso, movilidad del trabajador, etc.) se relacionan sólo marginalmente con las diferencias existentes entre las exposiciones de distintos trabajadores que ocupan puestos "aparentemente" equivalentes. Los esquemas basados únicamente en la "observación" deberían pues complementarse con una utilización juiciosa de los resultados de tomas de muestras efectuadas en trabajadores seleccionados aleatoriamente, tal como recomienda el borrador final de la norma europea EN 689(3).

La variabilidad observada entre los resultados obtenidos a partir de muestreos efectuados en un mismo puesto de trabajo en momentos distintos, es también elevada. Actualmente se considera generalmente admitido que dichos resultados se distribuyen según una ley de probabilidad logarítmico normal (4), si bien pueden presentarse excepciones. Por ello es prudente verificar la hipótesis de lognormalidad empleando una prueba estadística para verificar el ajuste. En la referencia (5), Waters ha recogido varias de dichas pruebas.

Una de las condiciones importantes para que se cumpla la hipótesis de lognormalidad de los resultados de las tomas de muestra ambientales es que éstas sean de duración **aproximadamente igual**. En efecto, (2), distintos estudios teóricos y experimentales han mostrado que muestras de distinta duración tomadas en el mismo ambiente de trabajo siguen distribuciones lognormales de la misma media pero de desviación standard tanto mayor cuanto más corto sea el tiempo de muestreo. En todo lo que sigue, pues, nos referiremos exclusivamente a los resultados relativos a tomas de muestra de duración sensiblemente idéntica.

Cuando una variable aleatoria, en este caso los resultados de las mediciones ambientales, sigue una ley de distribución logarítmico-normal, ello implica, por definición, que los logaritmos de dicha variable siguen una distribución normal o de Gauss. Ambas distribuciones se caracterizan mediante parámetros relacionados por expresiones matemáticas sencillas.

A efectos prácticos, los dos parámetros de mayor interés son la media aritmética y la desviación standard de las concentraciones μ_c y σ_c respectivamente. Como medida de dispersión es usual emplear, en lugar de la desviación standard, σ_c, el valor de la desviación standard geométrica, GSD (Geometric Standard Deviation) cuyos valores numéricos son más fáciles de manejar pues oscilan entre 1 y 5, aproximadamente. La GSD se define como:

$$GSD = \exp(\sigma_L)$$

donde $\exp(\sigma_L)$ significa e^{σ_L} y σ_L,es la desviación standard de la distribución de los logaritmos naturales de las concentraciones.

Incluso para valores de GSD relativamente bajos, la variabilidad "natural" de la distribución lognormal es considerable, lo que explica la dispersión habitualmente hallada en los resultados de las tomas de muestra ambientales. Para ilustrar cuantitativamente este hecho, en la tabla que se inserta a continuación (6) se indica, en el supuesto de que la media aritmética real de la concentración fuese 10 ppm, la amplitud del intervalo en el que se encontrarían el 50% de las muestras obtenidas para distintos valores de GSD; ello implica, naturalmente, que el 50% restante se encontrarían fuera de dicho intervalo.

GSD	INTERVALO, ppm
1,25	8,6 - 11,6
1,50	7,6 - 13,2
1,75	6,8 - 14,6
2,00	6,3 - 16,0
2,25	5,8 - 17,3
2,50	5,4 - 18,6

Puesto que los valores límite de la concentración ambiental de los contaminantes suelen estar definidos como medias aritméticas de la concentración durante un cierto periodo, uno de los objetivos importantes de toda medición ambiental consiste en determinar, con un margen de error conocido, la media aritmética de la distribución de las concentraciones.

Ello plantea dos principales problemas a resolver: la estimación, a partir de los resultados experimentales, del valor más probable de la media aritmética de la distribución de concentraciones y la determinación de un intervalo de confianza en cuyo interior se encuentre, con una probabilidad conocida (nivel de confianza), el valor real de dicha media aritmética.

Estimación de la media

El problema de estimar el valor de la media aritmética de la concentración a partir de los valores de las mediciones efectuadas fue tratado originalmente por Bar Shalom (7) y, más recientemente, por Selvin y Rappaport (8) y Attfield y Hewett (9), quienes, dicho sea de paso, a nuestro juicio no añaden nada nuevo a lo dicho hace casi treinta años por Bar Shalom.

Desde un punto de vista práctico, el problema puede enfocarse del siguiente modo: sean

c_1, c_2 ... c_n, un grupo de medidas independientes (véase el apartado "**Dimensión temporal del muestreo: el problema de la autocorrelación**" de esta NTP) de duración aproximadamente igual, l_1, l_2, ... l_n, sus respectivos logaritmos naturales y m_L la media de estos últimos:

$$m_L = \frac{\left(\sum l_i\right)}{n} \qquad \text{(a)}$$

La estimación μ^*_c de la media aritmética de la distribución (supuesta lognormal) de la que se han tomado las muestras, viene dada por la expresión:

$$\mu^*_c = (g_c)\varnothing(n, s^*_L) \qquad \text{(b)}$$

donde g_c es la media geométrica de las concentraciones de la muestra que, por definición, vale:

$$g_c = \exp(m_L)$$

y s^*_L es la estimación de la desviación standard σ_L de la ley de distribución normal que siguen los logaritmos de las concentraciones, que se calcula mediante la expresión:

$$s^*_L = \sqrt{\frac{\sum (m_L - l_i)}{(n-1)}} \qquad \text{(c)}$$

La función $\varnothing$ se expresa mediante un desarrollo en serie, por lo que su cálculo es engorroso. Por ello, en la figura 1se da su valor para distintos valores de n y GSD, habiendo tomado para el cálculo los cinco primeros términos de la serie, lo que da una aproximación suficiente (9). Quienes deseen evitar interpolaciones o, simplemente, obtener mayor exactitud, pueden recurrir a la expresión original:

$$\varnothing = 1 + \frac{n-1}{n} t + \frac{(n-1)^3}{n^2(n+1)} \frac{t^2}{2!} + \frac{(n-1)^5}{n^3(n+1)(n+3)} \frac{t^3}{3!} + \dots$$

donde:

$$t = \frac{(s^*_L)^2}{2}$$

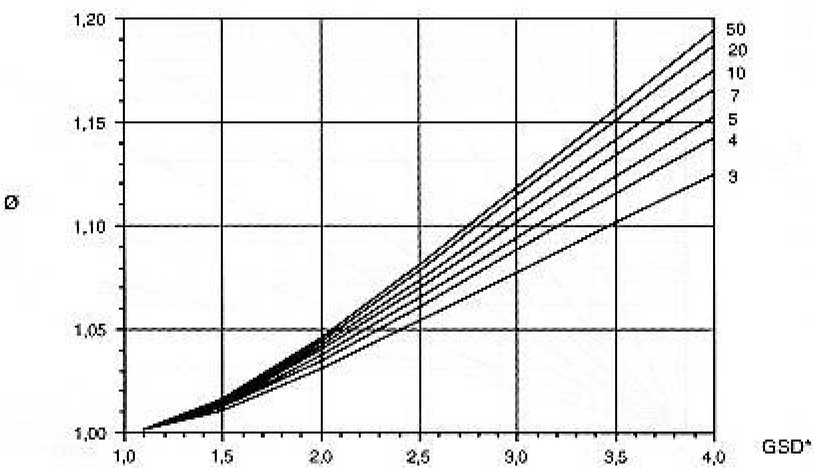

Fig. 1: Valores de ∅ en función de la estimación de la desviación standard geométrica GSD* y del número de muestras n

Ejemplo: Se han tomado cinco muestras de duración aproximadamente igual en momentos elegidos aleatoriamente (ver el apartado 5 de esta NTP). Los resultados obtenidos y sus respectivos logaritmos naturales fueron los siguientes:

c,mg/m^3	lnc(c)
2	0,693
3	1,099
4	1,386
5	1,609
6	1,792

A partir de los datos anteriores, la ecuación (a) da m_L = 1,316 y en consecuencia será g_c = 3,728; de la ecuación (c) se obtiene s^*_L = 0,434

Puesto que el valor obtenido de s^*_L es la estimación de s^*_L, la estimación GSD* de la GSD de la distribución de concentraciones, valdrá:

$$GSD^* = \exp(s^*_L) = 1{,}543$$

Con este valor, y teniendo en cuenta que el número de muestras es n = 5, la figura 1 da ∅ = 1,02.

En consecuencia, de la ecuación (b) se obtiene:

$$\mu^*_L = (g_c)\,\varnothing\,(n, s^*_L) = 3{,}728 \times 1{,}015 = 3{,}803$$

que es el valor más probable de la media aritmética de la concentración.

Intervalo de confianza de la media

La mejor estimación de la media, aun siendo un dato útil, es insuficiente para calibrar adecuadamente la situación. Tan importante como estimar la media es calcular su intervalo de confianza, que se define como aquel intervalo en el que existe una probabilidad conocida (nivel de confianza) de que se encuentre el verdadero valor de la variable.

Aunque los estudios fundamentales son relativamente antiguos (10), no ha sido hasta muy recientemente que se han publicado datos prácticos que permitan calcular con sencillez el intervalo de confianza de la media aritmética de la concentración a partir de los resultados de las mediciones ambientales (11).

El procedimiento operatorio es extremadamente sencillo; a partir de los valores c_1, c_2, ..., c_n se calculan, tal como se ha indicado, su media aritmética m_L y la estimación de la desviación standard s^*_L. A partir de estos valores se calcula la media geométrica gc de los valores c_1, ... c_n y la estimación de la desviación standard geométrica (GSD*):

$$gc = \exp(m_L)$$

$$GSD^* = \exp (s^*_L)$$

En la figura 2 se representa, en función de GSD* y para distintos valores del número de muestras tomado, n, el valor del factor de corrección F_{SUP} por el que hay que multiplicar el valor de g_c obtenido para obtener el extremo superior del intervalo de confianza de la media aritmética con un nivel de confianza del 95%.

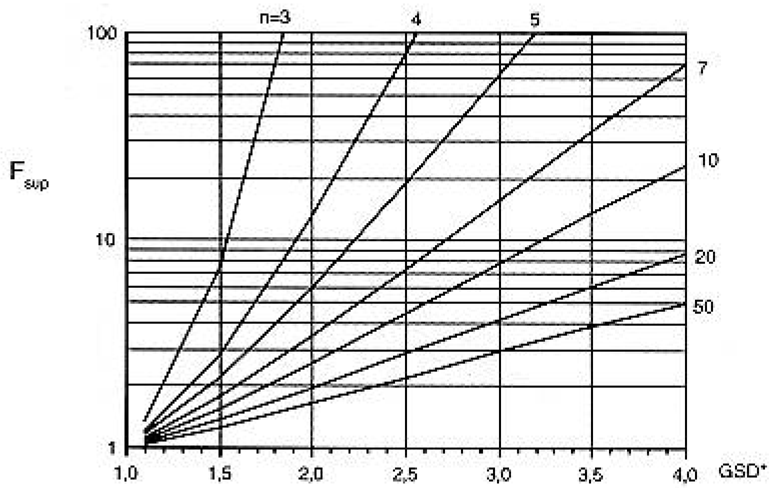

Fig. 2: Valores del factor de corrección F sup para el cálculo del extremo superior del intervalo de confianza de la media (nivel de confianza, 95%)

Análogamente, la figura 3 da los valores del factor de corrección F_{INF} por el que hay que multiplicar el valor de g_c para obtener el extremo inferior del intervalo de confianza de la media aritmética.

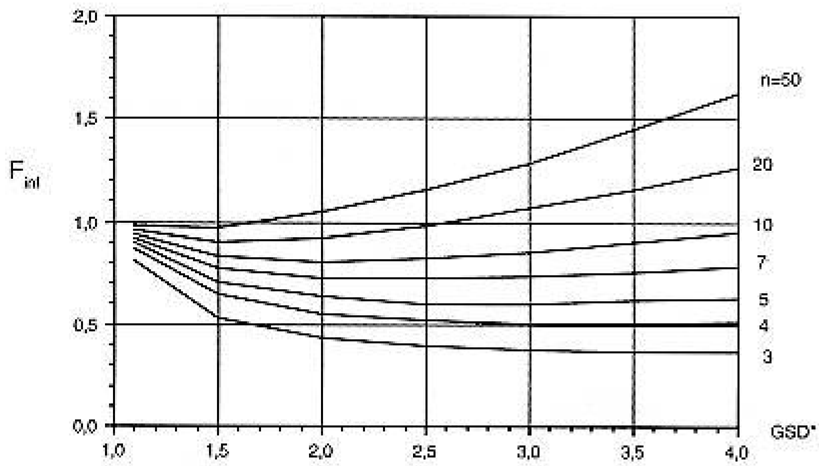

Fig. 3: Valores del factor de corrección F inf para el cálculo del extremo inferior del intervalo de confianza de la media (nivel de confianza, 95%)

Ejemplo: supongamos los valores de concentración del ejemplo anterior; tendremos:

$$m_L = 1{,}316 \quad g_c = 3{,}728 \quad GSD^* = 1{,}543$$

Entrando en la figura 2, con un valor de GSD = 1,543 y n = 5, da F_{inf} = 3,4; análogamente, empleando la figura 3 con los mismos valores de la desviación standard geométrica GSD y de n, se obtiene $F_{inf} = 0,7$.

Los extremos del intervalo de confianza, a un nivel de confianza del 95%, serán:

$$c_{sup} = g_c \times F_{sup} = 3,728 \times 3,4 = 12,675$$

$$c_{inf} = gc \times F_{inf} = 3,728 \times 0,7 = 2,610$$

Con los resultados experimentales obtenidos podemos afirmar que, con un nivel de confianza del 95% (es decir, en un 95% de los casos), el verdadero valor de la media se encontrará entre 2,610 y 12,675.

Dimensión temporal del muestreo: el problema de la autocorrelación

Si en un puesto de trabajo tomamos una muestra de, por ejemplo, quince minutos de duración y al cabo de varias horas o días repetimos la experiencia, la variabilidad «natural» de la distribución lognormal nos permite augurar una probabilidad elevada de que ambos resultados difieran apreciablemente.

En cambio, si las tomas de muestra se realizan de forma sucesiva, la intuición nos dice que, si no se han producido cambios detectables en el proceso, es poco probable que la diferencia entre ambos resultados sea importante. En otras palabras, el resultado de la primera medición "condiciona" el de la segunda si ambas se encuentran próximas en el tiempo: se dice que entre ambos resultados existe autocorrelación.

Los resultados experimentales confirman la intuición, si bien el grado de autocorrelación tiende a disminuir a medida que crece la duración de las muestras.

El fenómeno tiene un interés práctico notable, ya que si los resultados de las tomas de muestra efectuadas están correlacionados entre sí (autocorrelacionados), la validez de los cálculos que hemos visto en los párrafos anteriores (que se basan en la hipótesis de que los datos siguen la ley lognormal) se reduce notablemente.

Ello se debe a que los resultados de las mediciones ambientales sólo se distribuyen siguiendo una ley lognormal si son «independientes», es decir, si unos no condicionan a otros. En otras palabras, si existe autocorrelación no puede suponerse que los datos siguen una ley lognormal.

La existencia de autocorrelación puede introducir sesgos notables en los resultados de los cálculos que hemos visto en los apartados anteriores (que se basan en la hipótesis de que los datos siguen efectivamente una ley lognormal). Si el grado de autocorrelación es significativo, ello no sólo reduce la exactitud en la estimación de la media, sino que conduce a subestimar la magnitud de la desviación standard de la distribución, reduciendo así erróneamente la amplitud calculada del intervalo de confianza (el intervalo real es más amplio que el calculado).

El tema ha sido ampliamente debatido en los últimos años, ya que mientras algunos autores (12, 13) encuentran que el efecto es poco importante para muestras consecutivas y duraciones de tiempo de muestreo superiores a una hora, otros (14) llegan a una conclusión absolutamente opuesta, hallando altos niveles de autocorrelación incluso entre muestras de jornada completa tomadas en días consecutivos.

Mientras la balanza no se incline definitivamente en uno u otro sentido, se impone la prudencia y, por ello, se recomienda tomar muestras de duración lo más próxima posible a una jornada completa y seleccionar aleatoriamente las fechas de muestreo o, al menos, no efectuar las tomas de muestra en días consecutivos. El número de muestras a tomar dependerá de la amplitud del intervalo de confianza que se esté dispuesto a tolerar.

Bibliografía

(1) NICAS, M., SIMMONS, B.P., SPEAR, R.C
Environmental versus Analytical Variability in Exposure Measurements
Am. Ind. Hyg. Assoc. J. 52(12): 553-557 (1991)

(2) RAPPAPORT, S.M, KROMHOUT, H., SYMANSKI, E.
Variation of Exposure between Workers in Homogeneus Exposure Groups
Am. Ind. Hyg. Assoc. J. 54: 654-662 (1993)

(3) COMITÉ EUROPÉEN DE NORMALISATION
Workplace atmospheres: Guidance for the assessment of exposure by inhalation to chemical agents for comparison with limit values and measurement strategy
Final draft prEN689. April 1994

(4) RAPPAPORT, S.M.
Assessment of long-term exposures to toxic substances in air
Ann. Occup. Hyg. Vol 35. No. 1: 61-121 (1991)

(5) WATERS, A.M., SELVIN, S. RAPPAPORT, S.M.
A measure of Goodness-of-Fit for the Lognormal Model Applied to Occupational Exposures
Am. Ind. Hyg. Assoc. J. 52(11): 493-502 (1991)

(6) CASTELLÁ, J.L.
Estadística y mediciones ambientales. NTP-140
INSHT. Barcelona, 1985

(7) BAR-SHALOM, Y., SEGALL, A., BUDENAERS, D.
Decision and estimafion procedures for air contaminants
Am. Ind. Hyg. Assoc. J. 37(10):469-473 (1976)

(8) SELVIN, S., RAPPAPORT, S.M.
A Note on the Estimation of the Mean Value from a Lognormal Distribution
Am. Ind. Hyg. Assoc. J. 50(12): 627-630 (1989)

(9) ATTFIELD, M.D., HEWETT, P.
Exact Expressions for the Bias and Variance of Estimators of the Mean of a Lognormal Distribution
Am. Ind. Hyg. Assoc. J. 53(7): 432-435 (1992)

(10) LAND, C.E.
Standard Confidence Limits for Linear Functions of the Normal Mean and Variance J.
Am. Stat. Assoc. 68(344): 960-963 (1973)

(11) AMSTRONG, B.G.
Confidence Intervals for Arithmetic Means of Lognormally Distributed Exposures

Am.Ind. Hyg. Assoc. J. 53(8): 481-485 (1992)

(12) FRANCIS, M., SELVIN, S., SPEAR, R., RAPPAPORT, S.
The eff ect of Autocorrelation on the Estimation of Worker's Daily Exposures
Am. Ind. Hyg. Assoc. J. 50(1): 37-43 (1989)

(13) KUMAGAI, S., MATSUNAGA, I., KUSAKA, Y.
Autocorrelation of Short-Term and Daily Average Exposure Leveis in Workplaces
Am. Ind. Hyg. Assoc. J. 54(7): 341-350 (1993)

(14) BURINGH, E., LANTING, R.
Exposure Variability in the Workplace: Its Implications for the Assessment of Compliance
Am. Ind. Hyg. Assoc. J. 52(1): 6-13 (1991)