Bases estadísticas para establecer criterios de control ambiental mediante probabilidades de excedencia de concentraciones límites

ANTONIO RUIZ MATEO (*)

RESUMEN. En las II Jornadas Españolas de Ingeniería de Costas y Puertos se leyó una ponencia relacionada con los criterios de diseño de los vertidos al mar en la que se concluía que si en un punto de una zona de baño la concentración de coliformes fecales excedía un determinado valor c_{lim} durante una fracción p del tiempo, la probabilidad de que al tomar n muestras se obtuvieran más de np con concentración superior a c_{lim} era bastante alta y muy dependiente de n. De esta forma, para cumplir la norma de que no más de np muestras excedan la concentración c_{lim} era necesario diseñar el vertido de forma que el tiempo de excedencia p' fuera bastante menor que p.

En el presente artículo se deducen algunos de los resultados que se presentaron allí, se precisan conceptos e interpretaciones, se obtienen nuevos resultados y se dan bases para establecer criterios de control mediante probabilidades de excedencia.

STATISTICAL BASES FOR ESTABLISHING ENVIRONMENTAL CONTROL CRITERIA USING EXCEEDANCE PROBABILITIES OF LIMIT CONCENTRATIONS

ABSTRACT. *In the 2nd Spanish Symposium on Coastal and Harbour Engineering, a paper was read dealing with design criteria for waste disposal at sea, in which it was concluded that if the concentration of faecal coliforms in a bathing area exceeded a specific value c_{lim} for a time fraction p, it would be extremely likely that on taking n samples, more than np would contain a concentration greater than c_{lim} and this would be highly dependent on n. Thus, to fulfil the rule that with no more than np samples exceed the concentration c_{lim}, it was necessary to design the discharge in such a way that the exceedance time p' were considerably lower than p.*

This article deduces some of the results presented there, concepts and interpretations are determined, new results are obtained and the bases are laid for establishing control criteria through exceedance probabilities.

INTRODUCCION

En las II Jornadas Españolas de Ingeniería de Costas y Puertos, el profesor J. A. Revilla leyó una ponencia firmada por él y otros colaboradores [1] en la que se presentaban unos conceptos que a primera vista resultaban paradójicos. Se expusieron un conjunto de figuras en las que aparecía la probabilidad de que al tomar n muestras (con n variable representado en abcisas) de un punto en el que la concentración de coliformes fecales era superior a 2.000 org/100 ml durante una fracción p del tiempo (con $p = 0{,}05\text{-}0{,}02\text{-}0{,}01$), el número de muestras obtenidas con concentración superior a dicho valor límite era superior a np. Las dos conclusiones un tanto paradójicas que queremos destacar aquí son las siguientes:

a) Que si el tiempo de excedencia era de un 5 %, la probabilidad de que en un muestreo se obtengan más de un 5 % de las muestras era bastante alta y muy dependiente del tamaño de la muestra.

b) Que si se quiere cumplir con la norma impuesta por la Directiva 76/160/CEE relativa a la calidad de las aguas de baño [2], traspuesta a la legislación española por el Real Decreto 734/1988, de 1 de julio [3], de que no más del 5 % de las muestras tengan concentración superior a los 2.000 org/100 ml de coliformes fecales, el tiempo de excedencia de dichas concentraciones tiene que ser muy inferior al 5 %.

(*) Ingeniero de Caminos, Canales y Puertos. Jefe del Area de Ingeniería Marítima del Centro de Estudios de Puertos y Costas del CEDEX (Ministerio de Obras Públicas, Transportes y Medio Ambiente).

En el presente artículo se va a demostrar mediante la teoría de probabilidades que estas conclusiones, aunque paradójicas, son totalmente correctas. Además, se hará un planteamiento más general y se obtendrán nuevos resultados aplicables al control de cualquier parámetro de variación continua cuyo seguimiento se haga mediante muestreos discontinuos.

2. PLANTEAMIENTO DEL PROBLEMA

Llamaremos c al valor del parámetro de variación continua que se quiere controlar y que, en el ámbito de interés de este artículo, será normalmente la concentración de un contaminante.

Para el tratamiento estadístico que se va a hacer más adelante supondremos que el tiempo total T durante el cual se controla el parámetro c se divide en un número N de intervalos de tiempo ΔT elegidos de un tamaño tal que se cumplan las dos condiciones siguientes:

a) Que las concentraciones medias de los distintos intervalos de tiempo sean estadísticamente independientes.

b) Que el resultado obtenido al analizar una muestra sea representativo de la concentración media del intervalo durante el cual se tomó.

Llamaremos N_{sup} al número de intervalos con concentración media superior a un determinado valor límite c_{lim}. La fracción de tiempo de excedencia p vendrá dada entonces por el cociente:

$$p = N_{sup} / N \ (\leq 1) \qquad [1]$$

Planteado de esta forma, el problema es formalmente idéntico a otro muy conocido en teoría de probabilidades consistente en una caja con N bolas en su interior, de las cuales N_{sup} son de un color y el resto de otro. Analizar una muestra y comprobar si su concentración es superior a c_{lim} equivale a extraer una bola al azar y comprobar si es del primer color.

Esta analogía permite aprovechar la gran cantidad de resultados conocidos para el problema de las bolas sin más que hacer la adecuada interpretación en el ámbito del control ambiental.

PROBABILIDAD DE CONTROL NEGATIVO

El control del nivel de contaminación se hace tomando n muestras de n intervalos distintos de entre los N posibles. Diremos que el resultado del control es negativo cuando el número de muestras con concentración superior a c_{lim} es mayor que el producto np_{ref} siendo p_{ref} una probabilidad de excedencia adoptada como referencia en el criterio de control. Así, por ejemplo, la Directiva 76/160/CEE establece para los coliformes fecales $p_{ref} = 0{,}05$ con $c_{lim} = 2.000$ org/100 ml como valor imperativo y $p_{ref} = 0{,}20$ con $c_{lim} = 100$ org/100 ml como valor guía.

Mediante aplicación directa de la definición clásica de probabilidad como cociente de números de casos con $c > c_{lim}$ entre número de casos posibles se llega a que la probabilidad de que se obtengan ν muestras contaminadas es:

$$\pi_\nu^H = \frac{\binom{N_{sup}}{\nu} \cdot \binom{N - N_{sup}}{n-\nu}}{\binom{N}{n}} \quad (\nu \leq \min\,[n, N_{sup}]) \qquad [2]$$

que coincide con la conocida **distribución hipergeométrica.**

Por tanto, la probabilidad de que el control resulte negativo vendrá dada por la expresión:

$$P(\nu > np_{ref}) = \sum_{\nu > np_{ref}}^{n} \pi_\nu^H = 1 - \sum_{\nu = 0}^{\nu \leq np_{ref}} \pi_\nu^H \qquad [3]$$

Así, si se toma $N = 400$, la probabilidad de que el control resulte negativo para tiempos de excedencia del 5, el 3 y el 2 % toma los valores que se representan en la figura 1 en función del número n de muestras de control.

Este tipo de gráficos fueron los que se presentaron durante la lectura de la ponencia mencionada en la introducción y en ellos pueden comprobarse los comentarios que allí se hicieron.

Obsérvese que para $n = 1$ la probabilidad de que el control resulte negativo es exactamente igual a la fracción de tiempo de excedencia p.

APROXIMACIONES ASINTOTICAS

Con el fin de obtener unas expresiones más sencillas aplicables a determinados rangos de los parámetros y de facilitar los razonamientos que se harán más adelante, recordaremos aquí algunas aproximaciones asintóticas de la distribución hipergeométrica [4], [5], [6].

Puede demostrarse que la probabilidad π_ν^H dada por la expresión [2] está acotada de la siguiente forma:

$$\binom{n}{\nu}\left(p - \frac{\nu}{N}\right)^{\nu}\left(q - \frac{n-\nu}{N}\right)^{n-\nu} < \pi_\nu^H < \binom{n}{\nu} p^{\nu} q^{n-\nu}\left(1 - \frac{n}{N}\right)^{-n} \qquad [4]$$

siendo $p = N_{sup}/N$ y $q = 1 - p$.

Puesto que ambos extremos del intervalo tienden a coincidir cuando $n << N$, π_ν tendrá el mismo límite. Es fácil ver que dicho límite es:

$$\pi_\nu^B = \binom{n}{\nu} p^{\nu} q^{n-\nu} \qquad (\nu \leq n) \qquad [5]$$

que coincide con la **distribución binomial.**

A efectos prácticos puede suponerse que los valores de π_ν^H y π_ν^B coinciden para valores de N superiores a $10\,n$ [7].

Por ejemplo, en la tabla 1 aparecen las probabilidades de que el control resulte negativo para $p = 0{,}05$ y

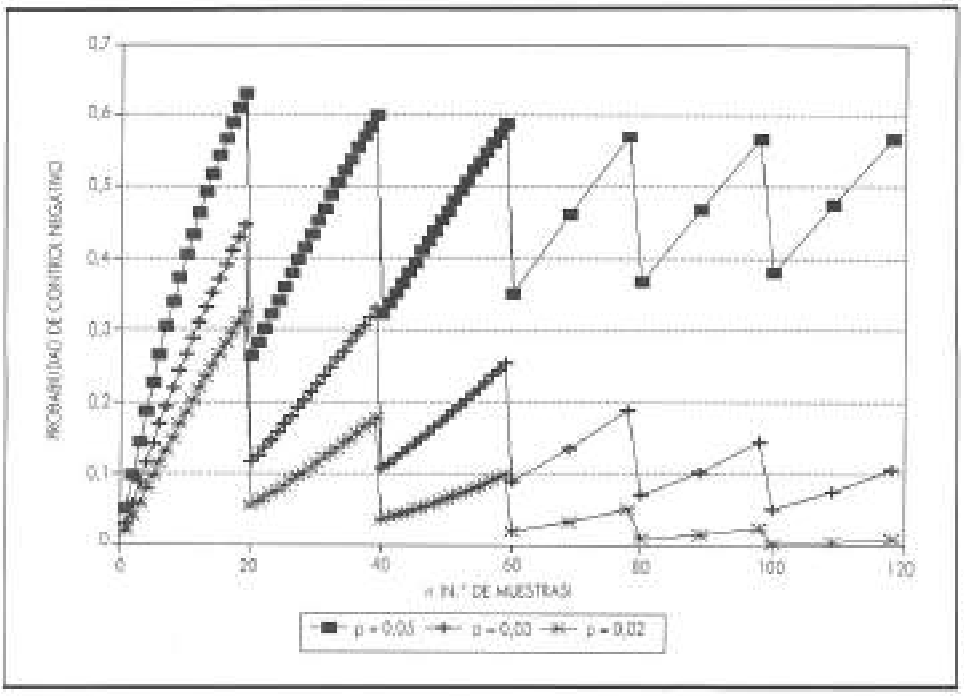

FIGURA 1. Probabilidad de que el control resulte negativo como función del tamaño de la muestra para distintos valores del tiempo de excedencia.

varios valores de n calculados a partir de la distribución hipergeométrica con $N = 400$ (las mismas que se representan en la figura 1) y las calculadas a partir de la distribución binomial.

Como puede verse, las diferencias son inferiores al 3 % en el caso más desfavorable ($n = 59$) e inferiores al 1 % en la mayoría de los casos.

Si pensamos que una duración razonable del intervalo de tiempo cuya concentración media es independiente de la del intervalo anterior puede ser de dos horas, el valor $N = 400$ representa un mes aproximadamente, lo que significa que si la duración T del período de control es superior a un mes, la distribución hipergeométrica puede sustituirse por la binomial, con la ventaja de que ésta queda definida por sólo dos parámetros (n y p) en lugar de los tres que se necesitan para determinar aquélla (n, p y N).

Veamos ahora la aproximación de Poisson. De [5] se tiene:

$$\pi_0^B(n, p) = q^n = (1-p)^n = \left(1 - \frac{np}{n}\right)^n \qquad [6]$$

Si suponemos que n es muy grande y p muy pequeño y llamamos λ al producto np, se puede hacer la siguiente aproximación:

$$\pi_0^B(n \cdot p) \simeq e^{-\lambda} \qquad [7]$$

Por otra parte:

$$\frac{\pi_\nu^B(n, p)}{\pi_{\nu-1}^B(n, p)} = \frac{\lambda - (\nu - 1)\,p}{\nu q} \simeq \frac{\lambda}{\nu} \qquad [8]$$

Aplicando esta expresión para $\nu = 1$ teniendo en cuenta [7] se obtiene $\pi_1^B(n, p) \simeq \lambda e^{-\lambda}/2$. Para $\nu = 2$ se obtiene $\pi_2^B(n, p) = \lambda^2 e^{-\lambda}/2$ y en general:

$$\pi_\nu^B(n, p) = \frac{\lambda^\nu}{\nu!} e^{-\lambda} \qquad [9]$$

La expresión:

$$\pi_\nu^P = \frac{\lambda^\nu}{\nu!} e^{-\lambda} \qquad [10]$$

representa la **distribución de Poisson**, que queda caracterizada por un solo parámetro ($\lambda = np$). Se considera [6] que ésta constituye una buena aproximación de la distribución binomial cuando se cumple simultáneamente que $n \geq 50$ y $p \leq 0{,}1$.

Como comprobación, en la figura 2 se representan los valores de ambas distribuciones para $n = 60$ y $p = 0{,}05$, siendo por tanto $\lambda = 3$. Puede verse que las diferencias son poco significativas y aún serán menores cuanto mayor sea n y menor sea p, como se deduce de la siguiente expresión que acota los valores de π_ν^B:

$$\pi_\nu^P(\lambda)\,\exp\left(-\frac{\nu^2}{n-\nu} - \frac{\lambda^2}{n-\lambda}\right) < \pi_\nu^B(n, p) <$$

$$< \pi_\nu^P(\lambda)\,\exp\left(\nu\,\frac{\lambda}{n}\right) \qquad [11]$$

La condición $n \geq 50$ se puede conseguir muestreando más de dos veces diarias si el período de control es un mes o muestreando más de dos veces por semana si el período de control es una temporada de baños de seis meses.

n	PROBABILIDAD		n	PROBABILIDAD		n	PROBABILIDAD	
	HIPERGEO	BINOMIAL		HIPERGEO	BINOMIAL		HIPERGEO	BINOMIAL
1	0,0500	0,0500	20	0,2640	0,2642	40	0,3220	0,3233
2	0,0980	0,0975	21	0,2830	0,2830	41	0,3370	0,3371
3	0,1430	0,1426	22	0,3020	0,3018	42	0,3520	0,3510
4	0,1860	0,1855	23	0,3220	0,3206	43	0,3660	0,3648
5	0,2270	0,2262	24	0,3410	0,3392	44	0,3810	0,3786
6	0,2660	0,2649	25	0,3600	0,3576	45	0,3950	0,3923
7	0,3040	0,3017	26	0,3790	0,3759	46	0,4100	0,4060
8	0,3390	0,3366	27	0,3980	0,3939	47	0,4240	0,4195
9	0,3730	0,3698	28	0,4160	0,4117	48	0,4390	0,4330
10	0,4050	0,4013	29	0,4340	0,4292	49	0,4530	0,4463
11	0,4350	0,4312	30	0,4520	0,4465	50	0,4670	0,4595
12	0,4640	0,4596	31	0,4700	0,4634	51	0,4810	0,4725
13	0,4920	0,4867	32	0,4880	0,4800	52	0,4950	0,4854
14	0,5180	0,5123	33	0,5050	0,4964	53	0,5090	0,4982
15	0,5430	0,5367	34	0,5520	0,5123	54	0,5230	0,5108
16	0,5670	0,5599	35	0,5380	0,5280	55	0,5360	0,5232
17	0,5890	0,5819	36	0,5540	0,5433	56	0,5490	0,5354
18	0,6110	0,6028	37	0,5700	0,5582	57	0,5620	0,5475
19	0,6310	0,6226	38	0,5850	0,5728	58	0,5750	0,5594
			39	0,6000	0,5871	59	0,5880	0,5711

TABLA 1. Comparación entre las probabilidades de control negativo con p = 0,05 calculadas a partir de la distribución hipergeométrica (N = 400) y de la distribución binomial.

La última de las aproximaciones asintóticas que vamos a recordar aquí es:

$$\pi_v^B(n, p) \approx \frac{1}{\sqrt{np(1-p)}}\,\varphi\left(\frac{v - np}{\sqrt{np(1-p)}}\right) \qquad [12]$$

o bien:

$$\pi_v^B(n, p) = \Phi\left(\frac{v + 0,5 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{v + 0,5 - np}{\sqrt{np(1-p)}}\right) \qquad [13]$$

siendo $\varphi(x)$ y $\Phi(x)$ las funciones de frecuencia y de distribución de la **distribución normal tipificada**:

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \qquad \Phi(x) = \int_{-\infty}^{x} \varphi(x)\,dx \qquad [14]$$

lo cual quiere decir que la distribución binomial puede aproximarse por una distribución normal de media $m = np$ y varianza $\sigma^2 = np(1 - p)$.

Esta aproximación es válida cuando se cumple que $np(1 - p) > 9$ y también vale como aproximación de la distribución de Poisson cuando $np > 10$.

Desde el punto de vista de su aplicación al problema que nos ocupa, se trata de una aproximación con más valor teórico que práctico porque tomando $p = 0,05$ resulta $n \geq 190$, lo que requería una muestra diaria durante un período de control superior a seis meses, condi-

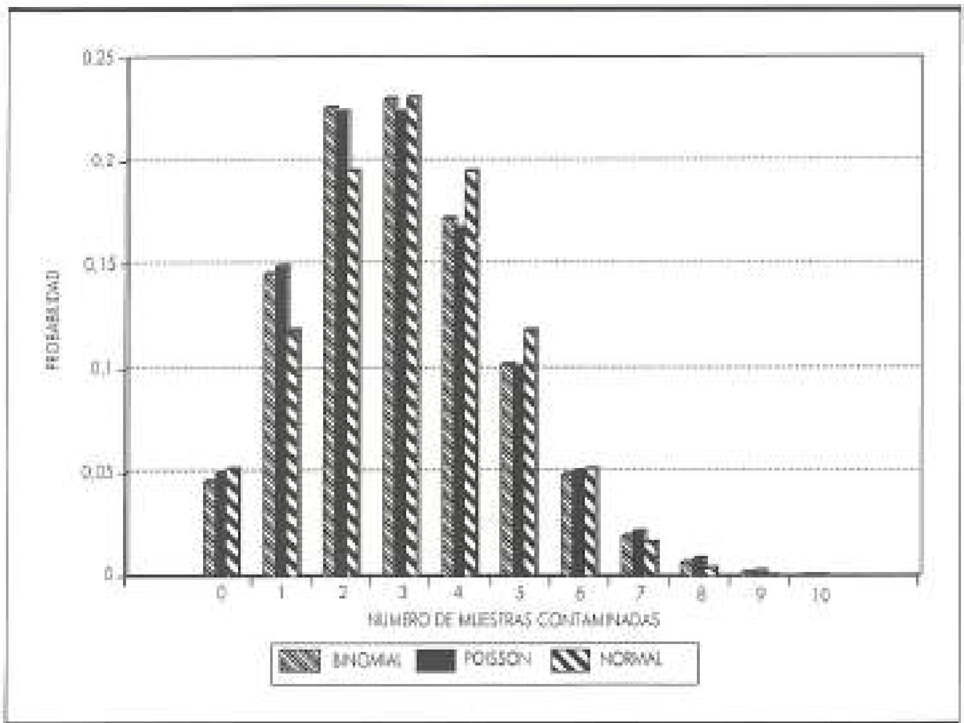

FIGURA 2. Comparación entre las distribuciones binomial, de Poisson y normal para n = 60, p = 0,05 y λ = np = 3.

ciones que están ya en el límite de las posibilidades prácticas. Sin embargo, puede resultar aplicable para el control de los valores guía de la Directiva 76/160/CEE, ya que se expresan con una probabilidad de excedencia $p_{ref}^{*} = 0{,}20$, lo que requiere sólo $n \geq 57$.

El valor teórico de esta aproximación se basa en la cantidad de propiedades conocidas de la distribución normal y se pondrá de manifiesto en el apartado siguiente.

En la figura 3 se representan los valores de las distribuciones binomial, de Poisson y normal para $n = 60$, $p = 0{,}2$ y $\lambda = 12$.

PROBABILIDAD DE CONTROL NEGATIVO CUANDO $p = p_{ref}$

En este apartado nos proponemos analizar teóricamente el siguiente problema: Si la fracción p de tiempo de excedencia de c_{lim} es exactamente igual a la probabilidad de excedencia p_{ref} adaptada como referencia en el criterio de control, ¿cuál es la probabilidad de que el resultado de control sea negativo?

A primera vista parece que debe ser nula puesto que los criterios de control expresados mediante una probabilidad de excedencia se establecen con la intención de que esa misma sea la fracción de tiempo de excedencia. Sin embargo, ya se ha comentado que los resultados de los cálculos (Fig. 1) dan que la probabilidad es bastante alta.

La explicación de esta paradoja la encontramos utilizando la aproximación [13], válida como sabemos para altos valores del producto np, o bien, si p viene dado, para altos valores de n. En efecto, la probabilidad de que el control resulte negativo es:

$$P(v > np_{ref}) = \sum_{v > np_{ref}}^{n} \pi_v^B(n, p_{ref}) \approx$$

$$\approx 1 - \Phi\left(\frac{np_{ref} + 0{,}5 - np_{ref}}{\sqrt{np_{ref}(1 - p_{ref})}}\right) = \qquad [15]$$

$$= 1 - \Phi\left(\frac{0{,}5}{\sqrt{np_{ref}(1 - p_{ref})}}\right)$$

Cuando n se hace muy grande, esta probabilidad tiende a $(1 - \Phi(0))$ y como la función de distribución normal tipificada es simétrica respecto a $x = 0$ se obtiene:

$$\lim_{n \to \infty} p(v > np_{ref}) = 0{,}5 \qquad [16]$$

Este importante resultado significa que **cuando $p = p_{ref}$, por mucho que aumentemos el número de muestras, la probabilidad de que el control resulte negativo no tiende a cero sino que tiende a ser del 50 %.** En efecto, del gráfico de la figura 1 se intuye ya este resultado.

La conclusión inmediata es que si se quiere imponer que la fracción del tiempo de excedencia de una cierta concentración límite no sobrepase un cierto valor p_{ref}, no debe establecerse como criterio que el número de muestras con concentración superior a dicho límite sea

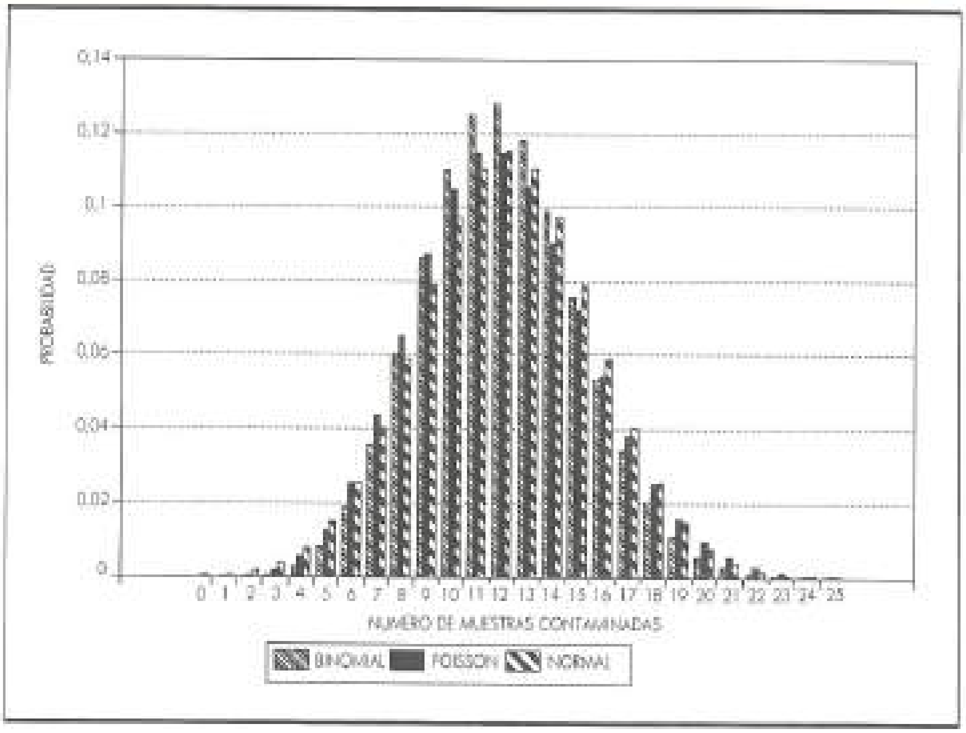

FIGURA 3. Comparación entre las distribuciones binomial, de Poisson y normal para $n = 60$, $p = 0{,}20$ y $\lambda = np = 12$.

inferior a np_{ref} porque entonces cuando el tiempo de excedencia cumpla estrictamente con el objetivo perseguido (es decir, que sea igual a p_{ref}) estarían resultando negativos aproximadamente la mitad de los controles.

Sería como si al imponer que un determinado tipo de monedas estuvieran perfectamente equilibradas se estableciera el criterio de que al arrojar n de ellas, el número de «caras» tuviera que ser inferior al 50 %. Aún estando equilibradas, de un conjunto de pruebas de n tiradas en cada una, por mucho que se aumentara el valor de n, la mitad de dicho conjunto daría como resultado que han salido más del 50 % de «caras».

La expresión [15] nos permite deducir además que el límite [16] se alcanza por debajo, es decir, que para grandes valores de n, los valores de la probabilidad p se mantienen inferiores a 0,5, ya que

$$\Phi\left(0{,}5 / \sqrt{np_{ref}(1-p_{ref})}\right) > 0{,}5$$

Por otra parte, se puede estimar la diferencia entre establecer como criterio que el número de muestras contaminadas sea inferior a np_{ref} o que sea no superior a np_{ref}. En efecto, la diferencia es exactamente el valor $\pi^{II}_{np_{ref}}$ que, de acuerdo con [12] y [14], puede estimarse mediante:

$$\pi^{II}_{np_{ref}} \simeq \frac{\varphi(0)}{\sqrt{np_{ref}(1-p_{ref})}} = \frac{1}{\sqrt{2\pi np_{ref}(1-p_{ref})}} \qquad [17]$$

Por ejemplo, para $n = 100$ y $p_{ref} = 0{,}05$ resulta $\pi^{II}_{np_{ref}} \approx 0{,}129$ que es un valor relativamente alto.

Con este criterio, de [13] y [15] puede obtenerse una expresión análoga a [16]:

$$\lim_{n\to\infty} p(\nu \geq np_{ref}) = 0{,}5 \qquad [18]$$

sólo que ahora el límite se alcanza por arriba.

Debemos recordar que estos resultados sólo son válidos si $np_{ref} > 10$. Para valores menores de n el comportamiento de $p(\nu \geq np_{ref})$ es el indicado en la figura 1.

BASES PARA ESTABLECER CRITERIOS DE CONTROL MEDIANTE PROBABILIDADES DE EXCEDENCIA

Retomando el símil del criterio para comprobar que las monedas están bien equilibradas, un criterio adecuado sería establecer unos límites inferior y superior de la razón frecuencial del número de «caras» determinados de forma que la probabilidad de que ésta se encuentre fuera de dichos límites sea inferior a una determinada α, lo que constituirá el riesgo de error del chequeo de control, es decir, la probabilidad de que se dé como no equilibrada una moneda que sí lo está. Como se sabe que al aumentar n la razón frecuencial se hace cada vez más próxima a 0,5, para un mismo riesgo de error el intervalo entre los límites será más pequeño cuanto mayor sea n, manteniéndose siempre el límite inferior por debajo de 0,5 y el límite superior por arriba. De esta forma, eligiendo los límites todo lo cercano que se quiera al valor 0,5, basta elegir n suficientemente grande para asegurar que la razón frecuencial se encontrará entre estos límites con una probabilidad tan alta como se quiera. Pero nunca puede alcanzar dichos límites el valor 0,5 porque entonces la probabilidad se haría exactamente cero.

Nuestro caso es un poco diferente porque no se trata de asegurar que la fracción de excedencia p sea exactamente igual a p_{ref}, sino de que no sea superior. Equivaldría a suponer que no importa que las monedas estén desequilibradas hacia el lado de las «caras» (de forma que salieran más «cruces»), pero que trataremos de imponer que no lo estén hacia el lado contrario.

La forma de acometer el problema es suponer que el objetivo de calidad se está cumpliendo estrictamente, es decir, que $p = p_{ref}$ (lo cual determina las características de la población) y **establecer un valor $p_{sup} > p_{ref}$ de forma que al tomar como criterio de control que el número de muestras contaminadas de un total de n muestras sea inferior a np_{sup}, la probabilidad de que el resultado sea negativo para la población con $p = p_{ref}$ esté por debajo de un valor pequeño α determinado, siendo α el riesgo de error del criterio adoptado.**

Por tanto, para que una norma de control ambiental del tipo que aquí se está tratando esté bien fundamentada sobre bases estadísticas debería incluir:

a) Una fracción p_{ref} del tiempo de excedencia de una determinada concentración límite c_{lim}, como elementos definitorios del objetivo de calidad.

b) Un valor α del riesgo de error con el que debe establecerse el control estadístico del cumplimiento del objetivo de calidad.

A partir de estos elementos, como se verá más adelante, ya es simple aplicación de las herramientas estadísticas la determinación de p_{sup} para cada valor de n. No obstante, la norma también podría fijar:

c) Un valor mínimo de n, para asegurar que la precisión del control estadístico está por encima de un determinado nivel.

La autoridad que se encargue del control deberá determinar p_{sup} para el valor elegido de n por encima de dicho mínimo. La expresión que define p_{sup} de forma implícita es:

$$\alpha = p\,(\nu > np_{sup}) = \sum_{\nu > np_{sup}}^{n} \pi_\nu^H (N, n, p_{ref}) \qquad [19]$$

o bien, cuando puedan utilizarse las expresiones asintóticas:

$$\alpha = \sum_{\nu > np_{sup}}^{n} \pi_\nu^B (n, p_{ref}) \qquad [20]$$

$$\alpha = \sum_{\nu > np_{sup}}^{n} \pi_\nu^P (n, p_{ref}) \qquad [21]$$

$$\alpha = 1 - \Phi\left(\frac{np_{sup} + 0,5 - np_{ref}}{\sqrt{np_{ref}\,(1 - p_{ref})}}\right) \qquad [22]$$

Con estos nuevos conceptos el resultado que se obtuvo en el apartado anterior y que se resume en la expresión [16] puede expresarse diciendo que para valores grandes de n, si se toma $p_{sup} = p_{ref}$, el riesgo de error α es del 50 %. Por otra parte, si $p_{sup} > p_{ref}$, por pequeña que sea su diferencia, α tenderá a $1 - \Phi\,(+\infty) = 0$, o dicho de otra manera, cualquiera que sea α, al crecer n, p_{sup} va a ir tendiendo a p_{ref}, como puede comprobarse en la figura 1.

Debe tenerse en cuenta que en todas estas expresiones, como np_{sup} debe de ser entero, el valor de α es una función discreta de np_{sup}. Esto quiere decir que en las expresiones [19] a [22] el signo igual debe sustituirse por mayor o igual, y p_{sup} debe elegirse como el mínimo valor que hace que se cumpla la desigualdad.

En la tabla 2 se da un listado de los valores de n calculados a partir de la distribución binomial (expresión [20]) en función de n para $\alpha = 0,10$ y $\alpha = 0,05$.

Alternativamente, en lugar del criterio b) anterior, la norma podría dar directamente una tabla del máximo número de muestras contaminadas en función del tamaño n de la muestra. Esto es lo que se hace, por ejemplo, en la Directiva 91/271 CEE sobre el tratamiento de las aguas residuales urbanas [8], donde se impone un «número máximo de muestras no conformes» en función del número de muestras tomadas en un año (cuadro 3 de la Directiva) que coincide con los valores de la tabla 2 para $\alpha = 0,05$.

INTERPRETACIONES DE LOS CRITERIOS DE LA DIRECTIVA 76/160/CEE

Como ya se ha dicho, la Directiva 76/160/CEE establece como valor imperativo para los coliformes fecales que menos del 5 % de las muestras tengan una concentración superior a 2.000 org/100 ml.

Por otro lado dice que la frecuencia de muestreo debe ser como mínimo «bimensual» durante la temporada de baños, lo cual se interpreta como dos muestras al mes porque la interpretación de una muestra cada dos meses resultaría una frecuencia extremadamente pequeña.

Puesto que no define un objetivo de calidad en términos de tiempo de excedencia, falta el valor p_{ref}. En cambio, se da $p_{sup} = 0,05$, pero sin indicar el valor de n al que está asociado. Por último, tampoco se da el valor de α.

Una interpretación posible es que el valor dado de p_{sup} corresponde a un valor de n teóricamente infinito. En este caso sabemos que p_{sup} tiende a p_{ref} (aunque sin llegar nunca a ser exactamente igual), por lo que podría decirse que la norma establece indirectamente el tiempo de excedencia. Esta es la interpretación que se le da normalmente y quizás sea la más adecuada pero, en ese caso, la autoridad controladora debe elegir el valor de α y, en función de éste, para cada tamaño de muestra n, calcular p_{sup}, o mejor aún el número np_{sup} máximo de muestras que pueden resultar contaminadas sin que el

NUMERO TOTAL DE MUESTRAS TOMADAS (n)		npsup
PARA ALFA = 0,10	PARA ALFA = 0,05	
1 a 2	1	0
3 a 10	2 a 7	1
11 a 22	8 a 16	2
23 a 35	17 a 28	3
36 a 49	29 a 40	4
50 a 63	41 a 53	5
64 a 78	54 a 67	6
79 a 94	68 a 81	7
95 a 109	82 a 95	8
110 a 125	96 a 110	9
126 a 141	111 a 125	10
142 a 158	126 a 140	11
159 a 174	141 a 155	12
175 a 191	156 a 171	13
192 a 200	172 a 187	14
	188 a 200	15

TABLA 2. Número máximo de muestras (np_{sup}) que pueden resultar contaminadas de un total de *n* muestras sin que el resultado de control deba valorarse negativamente (fracción de tiempo de excedencia pref = 0,05).

resultado del control sea negativo. En este sentido, los valores de la tabla 2 pueden resultar de gran utilidad.

Otra interpretación posible es que se trata de un valor de p_{sup} asociado a un valor de n a fijar por la autoridad que ejerce el control, por supuesto, por encima del mínimo bimensual. Aún así, todavía debe fijarse el valor de α porque según se desprende de la tabla 2, en bastantes casos un mismo número de muestras contaminadas puede hacer que un control resulte negativo dependiendo del valor de α, al ser diferente el riesgo de error aceptado.

Según esta interpretación, haciendo cálculos se obtiene que para un tamaño de muestra $n = 50$ el criterio de la Directiva podría corresponder a un tiempo de excedencia del 2,2 % si se acepta un riesgo de error de un 10 % o a un tiempo de excedencia del 1,65 % con un riesgo de error de un 5 %, mientras que para $n = 100$ correspondería a un tiempo de excedencia del 2,7 % con un riesgo de error del 10 % a un tiempo del 2,2 % con un riesgo de error del 5 %.

La misma duda de interpretación surge respecto a otros criterios de la Directiva.

En resumen, los cuatro parámetros p_{ref}, α, n y p_{sup} están ligados por una de las relaciones [19] a [22] por lo que pueden elegirse arbitrariamente (dentro de sus rangos posibles) tres de ellos. La Directiva sólo fija uno (p_{ref} o p_{sup}) y un valor mínimo para otro (n) por lo que existen dos grados de indeterminación que deberán ser fijados por la autoridad a cargo de ejecutar el control.

CONCLUSIONES

1.º) Cuando se quiere controlar el nivel de contaminación de un medio midiendo un determinado parámetro de variación continua mediante muestreos discontinuos, la probabilidad de obtener un determinado número de muestras contaminadas sigue una distribución hipergeométrica.

2.º) Bajo ciertas hipótesis de tamaño de la muestra y de tiempo de excedencia de la concentración límite, se pueden utilizar como aproximaciones de la anterior las distribuciones binomial (la más frecuente), de Poisson y normal.

3.º) Si se toma como criterio para valorar negativamente el resultado de un control el hecho de que el porcentaje de muestras contaminadas sea igual o mayor que el porcentaje del tiempo de excedencia, en una situación de cumplimiento estricto del porcentaje del tiempo de excedencia resultarían negativos el 50 % de los controles incluso para grandes valores de n.

4.º) Se propone un método estadísticamente coherente para establecer normas de control del nivel de contaminación. Se basa en definir una concentración límite c_{lim}, una fracción p_{ref} del tiempo de excedencia de ésta, un riesgo de error α de que resulte negativo un control cuando el tiempo de excedencia es exactamente p_{ref} y, opcionalmente, un valor mínimo del número de muestras que hay que tomar durante el período de control. De estos datos se deduce estadísticamente para cada número de muestras n el máximo número np_{sup} de éstas que pueden resultar contaminadas sin que el control se valore negativamente (Tabla 2).

5.º) Finalmente se analizan varias interpretaciones posibles de los criterios impuestos por la Directiva 76/160/CEE para el control de los coliformes fecales haciéndose notar que existen dos grados de indeterminación que han de ser elegidos por la autoridad a cargo de ejercer el control.

BIBLIOGRAFIA

1. REVILLA CORTEZON, J. A.; NIKOLOV KOEV, K.; ALVAREZ DIAZ, C., y TEMPRANO GONZALEZ, J. (1993). «¿Exige contaminación nula la Directiva comunitaria 76/160/CEE relativa a la calidad de las aguas de baño?» Leída en las II Jornadas Españolas de Ingeniería de Costas y Puertos. Gijón, 10 y 11 de mayo.
2. «Directiva 76/160/CEE relativa a la calidad de las aguas de baño». *Diario Oficial de las Comunidades Europeas*, núm. L 31/1 (5 de febrero de 1976).
3. Real decreto 734/1988 de 1 de julio, por el que se establecen normas de calidad de las aguas de baño. *Boletín Oficial del Estado*, núm. 167 (13 de julio de 1988).
4. CRAMER, H. (1968). «Elementos de la teoría de probabilidades». Aguilar.
5. FELLER, W. (1960). «An Introduction to Probability Theory and Its Applications». Volumen I. John Wiley & Sons.
6. COMMISARIAT A L'ENERGIE ATOMIQUE (1969). «Methodes statistiques en chimie analytique». Dunod.
7. KOROLIUK, V. S. (1981). «Manual de la teoría de probabilidades y estadísticas matemáticas». Mir.
8. «Directiva 91/271/CEE relativa al tratamiento de las aguas residuales urbanas». Mayo, 1991.

FERIA INTEGRAL DEL MEDIO AMBIENTE

NATUR '94

La preocupación por la conservación y mejora del medio ambiente se configura hoy como un principio esencial.
Para mejorar la calidad de vida es imprescindible conservar los ecosistemas naturales y su biodiversidad.
Como aprovechar racionalmente los recursos y la adaptación de los sistemas de producción y consumo o la modificación de actitudes sociales son ejes esenciales sobre los que gira NATUR.
En SEVILLA, en el Palacio de Exposiciones y Congresos del 5 al 8 de Mayo usted tiene una cita con la Naturaleza.

Del 5 al 8 de Mayo

JUNTA DE ANDALUCIA
Consejería de Cultura y Medio Ambiente
Agencia de Medio Ambiente

PALACIO DE EXPOSICIONES Y CONGRESOS
(Sevilla - Este). Apartado de Correos 4016
Fax: 4675350 - Teléfono: 4675140 - 41080 Sevilla.

FIBES
SEVILLA

IBERTUBO

FABRICADOS:
- TUBERIA DE PRESION
- TUBERIA DE RIEGO
- TUBERIA DE SANEAMIENTO
- PLACAS ONDULADAS Y NERVADAS
- PLACAS PINTADAS
- ACCESORIOS Y MOLDEADOS

Polígono Industrial de Toledo

CAPACIDAD DE PRODUCCION: 75.000 Tn./año

IBERTUBO, S.A.

OFICINAS CENTRALES:
Corazón de María, 6
Teléfono: (91) 416 28 00
Fax: 519 39 54
Télex: 48528 IBT
28002 MADRID

DELEGACIONES Y DISTRIBUIDORES EN TODA ESPAÑA