

Análisis de las instalaciones de agua con riesgo de legionelosis en la Costa del Sol, Málaga, España

Recepción: Enero de 2008 / Aceptación: Marzo de 2008

⁽¹⁾Enrique Gea Izquierdo

RESUMEN

En España, durante la última década, ha aumentado el interés por el mantenimiento higiénico-preventivo para el control microbiológico en las instalaciones de agua. Con objeto de obtener una óptima calidad en los parámetros de las instalaciones de intercambio aire-agua, se han desarrollado protocolos de tratamiento y normativa al efecto. Diversas investigaciones han puesto de manifiesto la existencia del agente biológico *Legionella pneumophila*. El estudio aborda la aplicación de métodos de mantenimiento preventivo con objeto de controlar el desarrollo de la bacteria. Se investigó una muestra de las instalaciones de agua correspondientes a edificios públicos de la Costa del Sol, Málaga (España), mediante 83 variables, análisis cluster y de correspondencias. Los resultados incluyen un análisis detallado de la aplicación de métodos de control de la calidad del agua según los tipos de edificios. Se concluye con el grado de cumplimiento en el mantenimiento y la importancia según la tipología.

Palabras clave: Legionelosis, instalaciones de agua, mantenimiento higiénico-preventivo, métodos de control.

ANALYSIS OF THE LEGIONNAIRES' RISK WATER FACILITIES IN THE SUN COAST OF MÁLAGA, SPAIN

ABSTRACT

In Spain, during the last decade, it has increased the interest by the preventive- hygienic maintenance for the microbiological control in the water facilities. Some protocols of treatment and norms have been developed to obtain an optimal quality in the parameters of the air- water interchange facilities. Diverse investigations have shown the existence of *Legionella pneumophila* biological agent. The study approaches the application of preventive maintenance methods with object to control the bacterium development. A water facilities sample corresponding to public buildings from the Coast of the Sun (Málaga, Spain) was investigated, by means of 83 variables, cluster and correspondences analysis. The results include a detailed analysis of the water quality control application methods according to the types of buildings. It concludes with the maintenance fulfillment degree and the importance corresponding to each typology.

Keywords: Legionnaires' disease, water facilities, preventive- hygienic maintenance, control methods.

INTRODUCCIÓN

En 1976, los investigadores del Centers for Disease Control and Prevention identifican al agente *Legionella pneumophila* como el culpable en un brote masivo de neumonía que afectó a los participantes en la American Legion convention [7].

La legionelosis es una enfermedad causada por una bacteria, perteneciente a la familia *Legionellaceae*. Esta familia comprende un género, *Legionella*, con un total de 40 especies y más de 50 serogrupos [2].

Más de la mitad de las especies han estado implicadas en infección humana [4]; sin embargo, la causa más común de legionelosis es la *Legionella pneumophila* serogrupo 1 [5,6] y el serogrupo más frecuente en el ambiente [1].

La bacteria *Legionella pneumophila* es el principal agente etiológico de la enfermedad del legionario, una forma de neumonía lobar. Ubicua en medio ambientes acuáticos, la bacteria, Gram- negativo es un facultativo intracelular parásito de protozoos. La patogénesis de la legionelosis se debe a la habilidad de la bacteria de invadir y crecer entre los macrófagos alveolares, y se cree ampliamente que esta habilidad resulta de una adaptación a los nichos intracelulares en la naturaleza [3].

En el medio ambiente urbano, la enfermedad se asocia usualmente con diseños deficientes en los edificios, construcción inapropiada y mantenimiento negligente de las instalaciones de agua. En los últimos cinco años han acontecido una serie de brotes en la provincia de Málaga (España), con especial repercusión en el turismo, siendo el más reciente el de junio de 2007, en la población de Benalmádena y fuente de origen en una torre de refrigeración del "Club Municipal de Hielo".

(1) Ph.D. Master en Ciencias. Profesor de Organización de Empresas, Universidad de Málaga (España).
 E-mail: enriquegea@telefonica.net

MATERIAL Y MÉTODO

Mediante un cuestionario de validación del estado de las instalaciones de agua y entrevista al “Director o Gerente” de cada edificio público, se recogieron datos relativos a 83 variables e interpretación sobre el porcentaje de existencia de tratamientos del agua de las instalaciones con riesgo de legionelosis. Las variables se “agruparon en 14 clusters/ tipos de instalaciones (I-XIV)”: agua caliente sanitaria/ acumuladores (I), agua fría sanitaria/ depósitos de agua (aljibes) (II), torres de refrigeración/ condensadores evaporativos/ condensadores adiabáticos (III), bañeras (sin recirculación, uso individual) (IV), piscinas (con recirculación, uso colectivo, spas o similar) (V), humectadores/ enfriamiento evaporativo/ otros (VI), líneas de agua en unidades dentales (VII), fuentes ornamentales (VIII), humectación en conductos de aire (IX), riego por aspersión (inclusive aerosol) (X), instalaciones termales (XI), acuarios (descubiertos) (XII), sistemas de agua contra incendios (XIII) y equipos de terapia respiratoria (XIV).

Se trata de un estudio descriptivo realizado sobre una muestra correspondiente a 214 edificios de uso público, análisis cluster (instalaciones) y de correspondencias.

El Análisis de Correspondencias se utiliza como un método estadístico para representar posibles asociaciones existentes entre variables categóricas o, mejor dicho, entre las clases que representan estas variables categóricas, con el objeto de establecer, si es posible, patrones o estructuras en los datos obtenidos con su observación. Se hace referencia al Análisis de Correspondencias asociado, al caso de dos variables “A: tipo de edificio” y “B: mantenimiento preventivo de las instalaciones” y a un test de independencia de caracteres de la χ^2 , consistente en descomponer el estadístico χ^2 en distintos componentes.

En el Análisis de Correspondencias se han de determinar coordenadas que permitan representar las clases (“valores”) de las variables en estudio, de forma que los valores más cercanos sean los más relacionados.

Primero, se deben transformar los datos en porcentajes para establecer una homogeneización de los mismos.

Los programas informáticos utilizados en la recopilación e interpretación de datos son, respectivamente: SPSS (Copyright SPSS Inc., 1989 - 2004. Windows (Version 13.0.1. 20 Nov 2004) y R Devel-

opment Core Team (2007), R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria.

RESULTADOS

Los resultados, por tipos de edificios analizados, son los del Cuadro N.º 1. Este anticipa las posibles relaciones entre los tipos de edificios y el grado de mantenimiento. La relación entre el tipo de instalación de agua y el número de edificios se muestra en el Cuadro N.º 2.

Tipo de edificio	Porcentaje
Centros de Salud	0
Administración (no Junta de Andalucía)	0.50
Junta de Andalucía	0.33
Hotel 2 estrellas	0.38
Hotel 3 estrellas	0.34
Hotel 4 estrellas	0.25
Hotel 5 estrellas	0.50
Hotel Gran Lujo	0.25
Instituto de Educación Secundaria	0.21
Centro Público de Educación Infantil y Primaria	0.51
Centro Privado de Formación Profesional (FP)	0.58
Centro Privado de Bachillerato	0.50
Centro Privado de Educación Infantil y Primaria	0.28
Centro Privado de Educación Secundaria Obligatoria	0.28

Fuente: Elaboración propia.

Estos apuntes se aprecian mejor si se representan los porcentajes en un eje de coordenadas (Figura N.º 1), en donde se presentan con mayor claridad las *distancias* entre los tipos de edificios consideradas respecto a la variable en estudio (grado de mantenimiento).

De esta forma, solo se observa una variable y se considera una *distancia* no muy relacionada con nin-

Cuadro N.º 2. Relación entre el tipo de instalación de agua/ número de casos (edificios).

Asociación de variables (Agrupación)	Número de casos totales
I	147
II	199
III	38
IV	17
V	16
VI	2
VII	0
VIII	13
IX	4
X	22
XI	1
XII	3
XIII	25
XIV	5

Fuente: Elaboración propia.

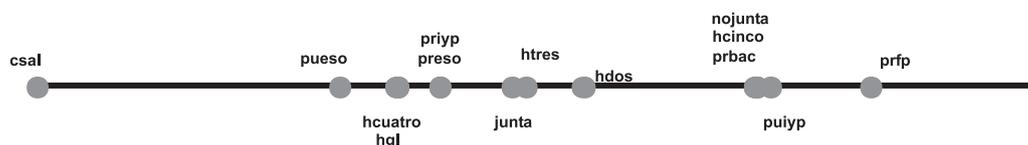
guna medida de asociación. Sería más razonable considerar como *distancia* la proporcionada por el estadístico λ de Pearson utilizado en el test de independencia de caracteres.

El estudio se realiza sobre $n = 214$ edificios clasificados en una tabla de doble entrada según la variable “tipos de edificios” que presenta los “valores” descritos, los cuales se han codificado respectivamente con sus “acrónimos” *csal*, *nojunta*, *junta*, *hdos*, *htres*, *hcuatro*, *hcinco*, *hgl*, *pueso*, *puiyp*, *prfp*, *prbac*, *priyp*, *preso*; y la variable “grado de mantenimiento higiénico-preventivo”, que presenta los “valores”: *cumplimiento*, *no cumplimiento* y *edificio sin instalaciones de riesgo*, codificados con los “acrónimos” *c*, *nc* y *si*.

En el contraste (test de la χ^2) de la independencia de las dos variables A y B, que presentan *a* y *b* clases respectivamente, en las que se clasifican las *n* observaciones; estas son recuentos del número de datos que aparecen en cada celdilla definida por una clase de cada variable y se presentan en forma de tabla de contingencia *a x b* según el Cuadro N.º 3.

A partir de los datos descritos, cabe plantearse el análisis para ver si están relacionadas las dos variables, con objeto de ver si el tipo de edificio afecta al grado de mantenimiento. Para ello se realiza un test de independencia de caracteres de la χ^2 , en donde la hipótesis nula sea la independencia de ambas variables. La hipótesis nula de independencia entre ambas variables se contrasta mediante el

Figura N.º 1. Representación de la distancia respecto al porcentaje del grado de mantenimiento.



Cuadro N.º 3. Descripción del grado de mantenimiento higiénico- preventivo/ tipo de edificio.

Tipo de edificio	c	nc	si	ni.
Centros de Salud	0	4	0	4
Administración (no Junta de Andalucía)	3	3	0	6
Junta de Andalucía	1	2	2	5
Hotel 2 estrellas	8	13	0	21
Hotel 3 estrellas	8	15	0	23
Hotel 4 estrellas	5	15	0	20
Hotel 5 estrellas	2	2	0	4
Hotel Gran Lujo	1	3	0	4
Instituto de Educación Secundaria (público)	5	18	0	23
Centro Público de Educación Infantil y Primaria	28	26	0	54
Centro Privado de FP	7	5	0	12
Centro Privado de Bachillerato	5	5	0	10
Centro Privado de Educación Infantil y Primaria	4	10	0	14
Centro Privado de Educación Secundaria Obligatoria	4	10	0	14
<i>n.j</i>	81	131	2	214

Fuente: Elaboración propia.

estadístico λ de Pearson, el cual mide las discrepancias existentes, para cada celdilla (i, j) , entre las frecuencias observadas n_{ij} en dicha celdilla y las frecuencias esperadas en ella, n_{ij}^t , si fuera cierta la hipótesis nula de independencia entre ambas variables, hipótesis que llevaría a que la probabilidad de obtener una observación en la celdilla (i, j) fuera $P(A_i, B_j) = P(A_i) \cdot P(B_j)$ y que en una muestra aleatoria simple de tamaño n , el número de observaciones que se obtuviesen en dicha celdilla sería una variable aleatoria binomial de parámetros $(n, P(A_i) \cdot P(B_j))$, siendo el número esperado de observaciones en dicha celdilla: $n_{ij}^t = n \cdot P(A_i) \cdot P(B_j)$.

Las probabilidades $P(A_i)$ y $P(B_j)$ no son conocidas y se estiman mediante sus estimadores máximo- verosímiles, que son las frecuencias relativas marginales, quedando el número esperado de observaciones en dicha celdilla igual a, $n_{ij}^t = n_i \cdot n_j / n$ y, finalmente, el estadístico λ de Pearson:

$$\lambda = \sum_{i=1}^a \sum_{j=1}^b \frac{(n_{ij} - n_i \cdot n_j / n)^2}{n_i \cdot n_j / n}$$

el cual, si la hipótesis nula de independencia es cierta y el tamaño muestral es suficientemente grande, sigue, aproximadamente, una distribución χ^2 con grados de libertad igual al número de celdillas de la tabla de contingencia en donde aparecen los recuentos de las n observaciones, menos uno, menos el número de parámetros estimados. Ya que solo hay que estimar $a - 1$ de las $P(A_i)$ y $b - 1$ de las $P(B_j)$ por sumar 1 estas probabilidades margina-

les, el estadístico λ de Pearson sigue una distribución χ^2 con $a \times b - 1 - [(a - 1) + (b - 1)] = (a - 1) (b - 1)$ grados de libertad.

Se obtiene el test Chi- cuadrado de Pearson = 100,1532, df = 26, p-valor = 1,210e-10 y las frecuencias observadas y esperadas (Cuadro N.º 4).

Para la contribución de cada celdilla al estadístico λ de Pearson se obtiene el valor del estadístico de Pearson, sumando los cuadrados de los valores de las celdillas y se calcula la matriz de discrepancias original,

	c	nc	si
csal	-1,23045467	0,99143798	-0,1933473
nojunta	0,48372616	-0,35111092	-0,2368011
junta	-0,64878226	-0,60631501	9,0358580
hdos	0,01823197	0,04040262	-0,4430143
htres	-0,23914628	0,24533515	-0,4636305
hcuatro	-0,93411059	0,78794242	-0,4323377
hcinco	0,39496076	-0,28668086	-0,1933473
hgl	-0,41774696	0,35237856	-0,1933473
pueso	-1,25591392	1,04485376	-0,4636305
puiyp	1,67236951	-1,22726269	-0,7104033
prfp	1,15330912	-0,86550695	-0,3348873
prbac	0,62448779	-0,45328224	-0,3057089
priyp	-0,56432772	0,48844411	-0,3617197
preso	-0,56432772	0,48844411	-0,3617197

El objeto es determinar coordenadas para cada una de las clases o "valores" de las dos variables que forman la tabla de doble entrada, con el propósito de representarlas. La determinación de las coordenadas implica una distancia entre los "valores". Se

Observadas	Esperadas		
	c	nc	si
csal	0	4	0
nojunta	3	3	0
junta	1	2	2
hdos	8	13	0
htres	8	15	0
hcuatro	5	15	0
hcinco	2	2	0
hgl	1	3	0
pueso	5	18	0
puiyp	28	26	0
prfp	7	5	0
prbac	5	5	0
priyp	4	10	0
preso	4	10	0

Fuente: Elaboración propia.

utiliza la denominada *distancia* Chi-cuadrado, dado que descomponemos el estadístico de Pearson.

Se determinan dos coordenadas para cada “valor”, de forma que estos se representan en un sistema cartesiano de ejes de coordenadas. En la determinación de las coordenadas de los “valores” de las variables es sumamente importante descomponer la matriz de discrepancias entre los valores observados y los valores esperados, señalado como C, en el producto de tres matrices, $C = UDV^t$; en donde las columnas de U corresponden con los *a* autovectores de CC^t ; las columnas de V corresponden con los primeros *a* autovectores de C^tC y D una matriz diagonal (de dimensión *a* x *a*); en donde los elementos de la diagonal son las raíces cuadradas de los *a* autovalores de CC^t .

Se cambia el nombre para la matriz de discrepancias (C), se determina la matriz U y se calcula la matriz V considerando que se han cambiado unos autovectores de signo (se cambia el signo de V para que se verifique la descomposición de la matriz C) para que se cumpla la igualdad anterior, no afectando a su definición, porque los autovectores se definen como soluciones de una ecuación igualada a cero, por lo que no existe problema en considerar unos valores o ellos mismos cambiados de signo. Esto implica, en todo caso, que las ordenadas pueden tomar signo contrario, de forma que representaciones obtenidas con distinto software podrían resultar simétricas respecto al eje de abscisas (es decir, en unos gráficos pueden quedar por arriba y en otros por abajo del eje de abscisas), siendo estas posibles discrepancias irrelevantes para el objetivo de que se trata, que es el de analizar la proximidad entre los “valores” de las variables y estos quedarán todos en uno u otro lado del mencionado eje.

La matriz D se determina, en primer lugar, sacando la raíz cuadrada de los autovalores de CC^t , y luego situándolos en la diagonal de una matriz.

Por último, se comprueban las matrices obtenidas y que su producto (C) coincide con la matriz de discrepancias original.

Las coordenadas correspondientes a las filas de la tabla de contingencia se obtienen de las columnas de la matriz U. Se buscan dos coordenadas para las filas, obteniéndose de los dos primeros autovectores de esa misma matriz, CC^t , es decir, de las dos primeras columnas de U. De igual manera, las coordenadas correspondientes a las columnas de la tabla de contingencia, se obtienen del primer o primeros autovectores de la matriz C^tC ; es decir, de la primera o dos primeras columnas de la matriz V. En general, las coordenadas para las filas y columnas se calculan según el Cuadro N.º 5.

Cuadro N.º 5. Cálculo de coordenada	
Filas	Columnas
$X_{j1} = U_{j1} \cdot \sqrt{\frac{n}{n_i}}$	$Y_{j1} = V_{j1} \cdot \sqrt{\frac{n}{n_j}}$

Fuente: Elaboración propia.

Aplicando lo anterior, las primeras y segundas coordenadas para las filas y columnas serán las indicadas en el Cuadro N.º 6.

El cálculo, y especialmente la representación, de las coordenadas correspondientes a los “valores” de las variables con que se forma la tabla de contingencia, se pueden obtener usando dos funciones de *módulo*. Se utiliza la función *correspondencia* (*x, nf*), en donde *x* es la matriz de datos (con estructura *data.frame*), y en donde el segundo argumento (=1) se usa para fijar el número de coordenadas a determinar para cada “valor” de las variables. Para representar las coordenadas de dimensión 2, se usa una nueva función de *gráfico de correspondencias* (*x*), con la misma estructura de *x*. Se consideran los

[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]	[,8]	[,9]
[1,] 9,187584	0,000000	0,000000e+00						
[2,] 0,000000	3,967556	0,000000e+00						
[3,] 0,000000	0,000000	8,971416e-08	0,000000e+00	0,000000e+00	0,000000e+00	0,000000e+00	0,000000e+00	0,000000e+00
[4,] 0,000000	0,000000	0,000000e+00	2,495196e-08	0,000000e+00	0,000000e+00	0,000000e+00	0,000000e+00	0,000000e+00
[5,] 0,000000	0,000000	0,000000e+00	0,000000e+00	1,912902e-08	0,000000e+00	0,000000e+00	0,000000e+00	0,000000e+00
[6,] 0,000000	0,000000	0,000000e+00	0,000000e+00	0,000000e+00	9,011062e-09	0,000000e+00	0,000000e+00	0,000000e+00
[7,] 0,000000	0,000000	0,000000e+00	0,000000e+00	0,000000e+00	0,000000e+00	5,983032e-09	0,000000e+00	0,000000e+00
[8,] 0,000000	0,000000	0,000000e+00	0,000000e+00	0,000000e+00	0,000000e+00	0,000000e+00	4,320931e-09	0,000000e+00
[9,] 0,000000	0,000000	0,000000e+00	0,000000e+00	0,000000e+00	0,000000e+00	0,000000e+00	0,000000e+00	6,329912e-10

Cuadro N.º 6. Coordenadas unidimensionales			
Primera coordenada		Segunda coordenada	
Fila	Columna	Fila	Columna
X11 = 0,13159633	Y11 = 0,12050337	X12 = -2,91899550	Y12 = 1,2757171
X21 = 0,1617326	Y21 = 0,08264913	X22 = 0,89234489	Y22 = -0,7916798
X31 = -6,4650898	Y31 = -10,29390472	X32 = 0,05111954	Y32 = 0,1884824
X41 = 0,1545573		X42 = -0,01511711	
X51 = 0,1525607		X52 = -0,26762827	
X61 = 0,1466645		X62 = -1,01332531	
X71 = 0,1617327		X72 = 0,89234489	
X81 = 0,1466645		X82 = -1,01332531	
X91 = 0,1446991		X92 = -1,26189098	
X101 = 0,1628488		X102 = 1,03350564	
X111 = 0,1667553		X112 = 1,52756828	
X121 = 0,1617326		X122 = 0,89234489	
X131 = 0,1488171		X132 = -0,74108671	
X141 = 0,1488171		X142 = -0,74108671	

Fuente: Elaboración propia.

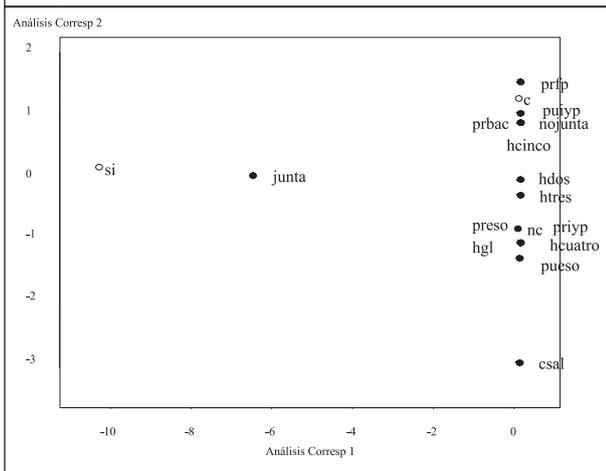
datos con la estructura mencionada y con posterioridad los nombres (“valores”) de las filas y las columnas. Para obtener una sola dimensión de las coordenadas asociadas a esos “valores” de las variables se ejecuta el *módulo correspondencia* (X), primera correlación canónica 0,6280503 y para la obtención de las coordenadas bidimensionales la correspondencia (X, nf = 2), primeras correlaciones canónicas 0,6280503 y 0,2712165.

Se obtienen los mismos valores (Cuadro N.º 7) que para el cálculo manual, sirviendo así de comprobación del proceso. La representación de las coordenadas (Figura N.º 2) se realiza por medio de un gráfico de correspondencias.

A continuación, se obtiene en un solo gráfico (Figura N.º 3) todas las representaciones de pares de componentes principales.

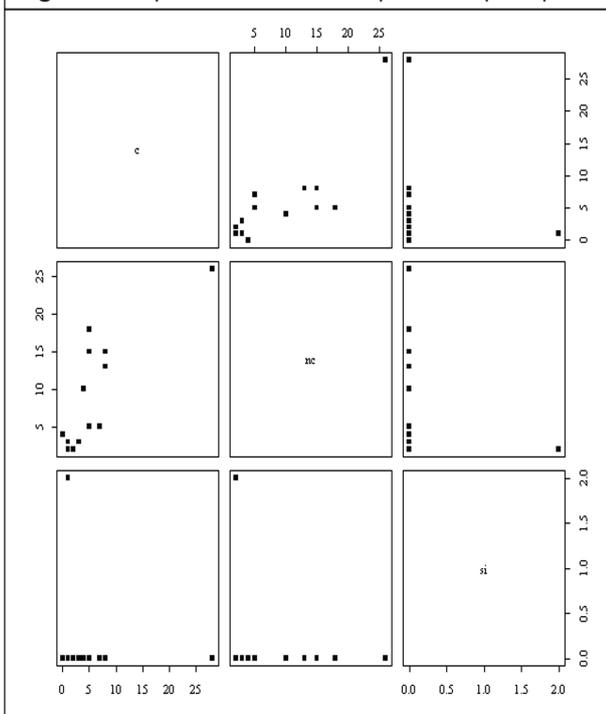
Cuadro 7. Coordenadas bidimensionales					
	Valores de las filas		Valores de las columnas		
	[,1]	[,2]	[,1]	[,2]	
csal	0,1315964	-2,91899550	c	0,12050337	1,2757171
nojunta	0,1617327	0,89234489	nc	0,08264913	-0,7916798
junta	-6,4650899	0,05111954	si	-10,29390472	0,1884824
hdos	0,1545574	-0,01511711			
htres	0,1525608	-0,26762827			
hcuatro	0,1466645	-1,01332531			
hcinco	0,1617327	0,89234489			
hgl	0,1466645	-1,01332531			
pueso	0,1446991	-1,26189098			
puiyp	0,1628488	1,03350564			
prfp	0,1667554	1,52756828			
prbac	0,1617327	0,89234489			
priyp	0,1488171	-0,74108671			
preso	0,1488171	-0,74108671			

Figura N.º 2. Representación de correspondencias



Fuente: Elaboración propia.

Figura 3. Representación de componentes principales



Fuente: Elaboración propia.

Para analizar si la utilización de las dos primeras coordenadas es adecuada, se aplica la *dimension* (x) obteniendo la: “Inercia total” (= 0,4680056), “Inercia con 1 dimensión” (0,3944472 - 84,2825800%), “Inercia con 2 dimensiones” (0,4680056 - 100,000000%) e “Inercia con 3 dimensiones” (0,4680056 - 100,000000%).

CONCLUSIONES

El Cuadro N.º 1 refleja que los centros de salud no tienen relación en cuanto a un mantenimiento adecuado; sin embargo, sí existe una relación entre institutos de Educación Secundaria, hoteles de 4 Estrellas, hoteles Gran Lujo, centros privados de Educación Infantil y Primaria y centros privados de Educación Secundaria Obligatoria. Asimismo, se pueden relacionar los edificios de la Junta de Andalucía, hoteles de 3 Estrellas y hoteles de 2 Estrellas, estableciendo una estrecha relación entre los dos primeros. Se produce una coincidencia en el porcentaje de los edificios de la Administración (no Junta de Andalucía), hoteles de 5 Estrellas y centros privados de Bachillerato, y casi de estos tres con los centros públicos de Educación Infantil y Primaria. Menos relacionados con los anteriores se encuentran los centros privados de FP.

Se observa un p-valor del test muy pequeño, por lo que se rechaza la hipótesis nula de independencia entre ambas variables. Con el Análisis de Correspondencias se detecta si es posible deducir que ciertos “tipos de edificio” tienen más posibilidades higiénico- preventivas o no o como afecta sobre el cumplimiento el grado de mantenimiento.

De la Figura N.º 2, se obtienen las principales conclusiones por la proximidad o lejanía de los “valores” de las variables. Al estar *si* y *junta* apartados de los demás se puede concluir que ambos no tienen relación con el resto, sí entre ellos. Por otra parte, sí se puede establecer una relación entre *c* y *prfp* y *puiyp* y menor con la “representación/ es” *hcinco- nojunta- prbac*. Existe una gran relación entre *nc* y *preso- priyp* y menor con *hgl- hcuatro, pueso, htres* y *hdos* (de mayor a menor relación). Se observa una relación entre *csal* y *nc*.

La tercera componente principal señala al dato tercero (*junta*) como anómalo, dato cuya simple observación en la matriz de datos hace sospechar su carácter. Los datos anómalos no son despreciables, ni algo a eliminar, sino que su identificación puede inclusive ser el objetivo que se persiga, no siendo el caso.

La primera componente principal, cruzada con las otras y especialmente con la segunda señala, aunque en mucha menor medida, al primer dato (*csal*) también como algo sospechoso por extremo, aunque dentro de la dirección de la componente principal.

Se ha presentado que se puede determinar una o dos (o inclusive más) coordenadas para cada "valor" de las dos variables con las que se realiza un Análisis de Correspondencias. Un Análisis con solo una coordenada puede conducir a resultados incompletos. Es necesario fijar cuántas coordenadas, y se deben determinar de una forma razonable, o al menos poder valorar que la determinación de dos coordenadas corresponde con una decisión acertada. Las componentes principales se aplican para la detección de valores anómalos en el conjunto de datos. Partiendo de esto, se puede sospechar que otras representaciones de pares de componentes principales pudieran desenmascarar otros datos anómalos.

Es conveniente definir cómo a la suma de todos los autovalores dividido por la frecuencia total n , se denomina *Inercia total* y, si los dos primeros autovalores (divididos por n) recogen ya la casi totalidad de la Inercia, la representación bidimensional es adecuada. Para ello y haciendo referencia al Análisis de Componentes Principales, se presenta el hecho de analizar el tamaño del primer o primeros autovalores en relación a la suma de todos ellos, disponiendo del primer o primeros autovectores si aquellos recogen la casi totalidad de la suma de autovalores. Los autovalores a analizar son los asociados a los autovectores de la matriz CC^t y, cada uno de ellos, se denomina *Inercia principal*. Se puede observar la primera dimensión en términos absolutos y en tanto por ciento, no siendo suficiente éste último. Asimismo, una representación en dos dimensiones recoge el 100%, suficiente como para concluir que es adecuada la representación bidimensional.

En el grado de cumplimiento del mantenimiento higiénico-preventivo frente a la legionelosis se puede concluir que es insuficiente en todos los tipos de edificios muestreados, existiendo un riesgo aparente y considerable de desarrollo de la bacteria.

REFERENCIAS BIBLIOGRÁFICAS

Bartlett, C.R.L., Kurtz, J.B., Hutchison, J.G., Turner, G.C. y Wright A.E. (1983). *Legionella* in hospital and hotel water supplies. *Lancet*. ii: 1315.

Benson, R.F., Thacker, W.L., Wilkinson, H.W., Fallon, R.J. y Brenner, D.J. (1988). *Legionella pneumophila* serogroup 14 isolated from patients with fatal pneumonia. *J. Clin. Microbiol.* 26 (2):382.

Cianciotto, N.P. (2001). Pathogenicity of *Legionella pneumophila*. *Int. J. Med. Microbiol. nov*; 291(5): 331-343.

Lo Presti, F., Riffard, S., Vandenesch, F., Reyrolle, M., Ronco, E., Ichai, P. y Etienne, J. (1997). The first clinical isolate of *Legionella parisiensis*, from a liver transplant patient with pneumonia. *J. Clin. Microbiol.* 35 (7): 1706 - 1709.

Marston, B.J., Lipman, H.B. y Breiman, R.F. (1994). Surveillance for legionnaires' disease: risk factors for morbidity and mortality. *Arch. Intern. Med.* 154: 2417 - 2422.

Reingold, A.L., Thomason, B.M., Brake, B.J., Thacker, L., Wilkinson, H.W. y Kuritsky, J.N. (1984). *Legionella* pneumonia in the United States: the distribution of serogroups and species causing human illness. *J. Infect. Dis.* 149: 819.

Swanson, M.S. y Hammer, B.K. (2000). *Legionella pneumophila* pathogenesis: a fateful journey from amoebae to macrophage. *Annu. Rev. Microbiol.* 54: 567-613.