

Descripción de dos métodos para la investigación causal en accidentabilidad laboral: El análisis de senderos y los dendogramas.

Miguel Verdeguer Cuesta

Licenciado en Psicología, Ingeniero Técnico Industrial.

Técnico de Prevención FREMAP.

RESUMEN

Desde siempre se ha considerado la importancia, en mayor o menor grado, del factor humano (centrado en el trabajador) y del factor técnico (máquina, equipo de trabajo, instalación.....), en cualquier accidente laboral.

Ambos factores, no son exclusivos sino complementarios, originándose el accidente en algunos casos por el predominio de un factor sobre el otro.

Cuando aún no existía una legislación y normativa efectiva en seguridad laboral y el diseño de las máquinas respondía tan solo a criterios de producción era frecuente en los accidentes de trabajo la relevancia del factor técnico sobre el humano.

Actualmente esto ya no es así, la Ley de Prevención de Riesgos Laborales en su artículo 41 obliga a los fabricantes, importadores y suministradores de maquinaria a asegurar y documentar que la maquinaria no constituye una fuente de peligro para el trabajador; ello reduce de forma notable la incidencia del factor técnico en los accidentes.

No vamos a negar que hay accidentes cuya causa material es una máquina, pero por regla general, el accidente ocurre no por un fallo o rotura de la misma (factor técnico) sino por un uso inadecuado o bien por la anulación de algún dispositivo de seguridad por parte del operario, acción que ejecuta no

con mala fe, sino con la idea de facilitar el trabajo desarrollado, acceder mejor al ajuste de la máquina, disminuir tiempos muertos empleados en retirar y colocar protectores y resguardos.....etc.

Si el accidente ha sido originado por un fallo, rotura o desgaste en una máquina o elemento de la misma, resulta relativamente sencillo identificar la causa y tomar las medidas oportunas; no ocurre lo mismo cuando el percance es imputable a un descuido humano (que normalmente no se quiere reconocer) al no seguir el método de trabajo indicado, manipular una máquina sin el aprendizaje adecuado...etc.

El que un trabajador en un momento determinado (estado) o bien de forma habitual (rasgo) no acate las normas y consignas de seguridad, las ignore o bien su motivación e interés no sea el adecuado en los temas relacionados con la seguridad y salud laboral, facilitará la posibilidad de ser víctima de un accidente o bien de provocarlo originando incluso daños a terceros.

Si, al igual que ocurre con el factor técnico, podemos con respecto al factor humano recoger una serie de variables tales como: Edad del accidentado, antigüedad en la empresa, hora y día de la semana en el que ocurrió el accidente, días de baja o duración del accidente, nivel aptitudinal, formación o destreza en la labor realizada, actitud frente a la seguridad, hora de ocurrencia...etc, quizás podamos de

alguna forma encontrar semejanzas o similitudes entre los accidentes. Evitando la acumulación de estas variables cabe la posibilidad de impedir o reducir los efectos del accidente.

En el punto que nos encontramos, se puede pensar que estamos propugnando por el establecimiento de un modelo determinista o defendiendo una « conexión necesaria » parafraseando a Hume entre causas-accidente o bien pretendiendo elaborar un modelo causa-efecto que responda a una « ley necesaria » (Kant) y no es así, mas bien nos inclinaríamos por intentar conocer o detectar relaciones y regularidades (Comte) entre los accidentes-accidentados lo cual nos llevaría a un mejor conocimiento del accidente y del entorno multicausal donde se materializa.

Así pues, nuestro objetivo no persigue obtener una función o ecuación (modelo matemático) sino observar la presencia de relaciones fuertes y no aleatorias entre una serie de variables obtenidas a partir, en este caso, de los datos reflejados en el parte de accidente.

Al existir similitudes y relaciones entre los accidentes o los sujetos que se accidentan, podemos inferir la no aleatoriedad del fenómeno accidente y actuar sobre las causas comunes reduciendo o impidiendo su materialización.

Si consultamos bibliografía y estudios al respecto, encontraremos estudios de tipo correlacional entre dos variables en los que los autores, mediante ajustes de tipo lineal o mediante una transformación logarítmica obtienen modelos que tientan a predecir una determinada variable dependiente sencillamente introduciendo datos en una ecuación.

Al respecto sería conveniente indicar que siempre que hagamos uso de la estadística, sobre todo de la multivariada, debemos tener presente que «Las estadísticas no son sustituto del juicio» (Henry Clay) y que las conclusiones que se obtienen no deben ser sacadas de contexto cito como ejemplo: «Los médicos hallan una asociación entre menor nivel de estudios y mayor tasa de obesidad» (Las Provincias 25/ 4/96). ó «La mortalidad por cáncer de útero es mayor en mujeres con baja formación» (Levante) o mejor aún: «Estar casado e ir a misa disminuye el riesgo de infarto» (Ag. Europa Press. IV congreso hipert. arterial. Huelva).

Por lo tanto, cuidado con las conclusiones que «extraigamos» de nuestro análisis estadístico.

El análisis de senderos o path-analysis es un método confirmatorio de tipo Durkheniano incluido dentro de las denominadas técnicas de análisis multivariante y fue fundamentado por S.Wright en 1921. El método parte de un modelo lineal y relaciona una serie de variables, indicando el grado de dependencia de unas sobre otras, para ello es necesario conocer la matriz de correlaciones entre todas las variables intervinientes y las ecuaciones estructurales que las ligan entre si.

El número de variables, si es elevado y estas discriminan no conlleva ningún problema, exceptuando el debido al manejo de matrices de dimensiones elevadas, cuestión que puede resolverse mediante una hoja de cálculo o un programa de aplicaciones estadísticas.

El análisis de senderos requiere, para

su correcta aplicación que se cumplan los siguientes supuestos:

a) Las variables dependientes deben estar completamente determinadas por una combinación

de las variables intervinientes.

b) Dos variables no pueden ser causa y efecto una de otra.

c) El modelo debe ser lineal o se deberán efectuar las transformaciones pertinentes para

que lo sea, de tipo logarítmico.....etc.

d) Las variables a utilizar deben ser continuas (de intervalo o de razón) de tipo cuantitativo .

En el caso de disponer de variables cualitativas, éstas deben ser previamente dicotomizadas aplicándoles el correspondiente coeficiente de correlación adecuado para su tratamiento estadístico.

cidas) entre ellos y la representación en un gráfico de estas relaciones vendrá dada por un gráfico a modo de árbol denominado dendograma. El dendograma forma grupos o clusters con aquellos casos-accidentes similares y a la vista del mismo es muy fácil observar si éstos están relacionados entre si.

Ejemplo (Caso real): Se trata de establecer las relaciones de dependencia y la matriz de similitudes entre los accidentes en una empresa del sector metal con una media de 50 operarios y un registro de 18 accidentes en un año.

La primera operación a realizar, será disponer de los 18 partes de accidente, impresos que nos servirán para localizar los siguientes datos-variables:

Figura I

TIPO DE VARIABLE	Coef. correlación a utilizar.
Las dos de razón	Pearson
Una de razón y la otra dicotómica	Biserial - Puntual
Las dos dicotómicas	Coeficiente «phi»

La otra técnica a utilizar para indagar las similitudes o relaciones entre los accidentes o los sujetos es el análisis de conglomerados, método desarrollado por Sneath y Sokal en 1963, el método se aplica sobre una matriz de medidas sobre n situaciones, la matriz es previamente transformada en otra de dimensiones nxn o mxm conocida como matriz de distancias o similitudes, donde cada uno de sus elementos representa la distancia euclidean (previamente normalizada) entre el y los otros elementos de la matriz, es decir se obtiene una matriz cuadrada cuya diagonal principal son todos ceros, pues la distancia o similitud entre un accidente y el mismo es nula, en el caso de que la distancia entre varios accidentes sea reducida o similar, ello implicará la similitud (causas pare-

- (X1): Mes de ocurrencia del accidente.
- (X2): Edad del accidentado.
- (X3): Día de la semana en el que ha ocurrido el accidente.
- (X4): Antigüedad en la empresa. Usar meses en vez de años.
- (X5): Duración del accidente en días. Días de baja.
- (X6): Número de horas trabajando hasta que ocurre el accidente. N Normalmente de 1 a 8 hrs.

Observaciones: El número de variables puede ser mayor; si previamente hemos pasado pruebas a los operarios y disponemos de referencias de su preparación para la realización del trabajo (nivel de aptitud y destrezas), actitud y motivación por los aspectos relacionados con la salud y seguridad laboral,

DOSSIER

podremos disponer de mayor información.

Hay que indicar también, que no siempre son exactos, rigurosos o fiables los datos extraídos del parte de accidente, así es fácil encontrar errores en cuanto al número de horas que el operario llevaba trabajando hasta el momento del accidente o en la hora de ocurrencia del mismo.

Para la elaboración del diagrama de senderos, deberemos elegir la variable dependiente, que en nuestro caso será el número de días que el operario ha estado de baja: (X5).

A continuación, dibujaremos un grafo en el que mediante flechas relacionaremos a todas las variable entre si, teniendo en cuenta que las flechas irán de las variables independientes a la/s

dependientes. A la vista del grafo (Que quizás posteriormente haya que redefinir) plantearemos las ecuaciones estructurales, con ellas y con los coeficientes de correlación entre las variables (todas) que previamente habremos calculado, estimaremos los coeficientes de Wright, mediante los cuales conoceremos el grado de dependencia o influencia de las variables entre si.

Si las relaciones de dependencia son bajas, habrá que modificar el modelo cambiando las relaciones entre las variable iniciales.

No obstante y previo al diseño del modelo, es conveniente obtener el dendograma o representación gráfica de las similitudes o distancias de los accidentes-accidentados, un distanciamiento elevado entre los accidentes y la nula obtención de clusters (grupos) indicará una cierta dificultad a la hora de obtener un modelo causal de los mismos.

Para la elaboración del modelo, unimos cada variable independiente con la dependiente mediante flechas, de tal modo que la flecha va de la variable independiente a la dependiente.

(Ver figura IV)

Planteamiento de las ecuaciones estructurales: (p= coeficientes deWright)

- 1) $X3 = p23 X2$
 - 2) $X4 = p43 X3 + p24 X2$
 - 3) $X5 = p54 X4 + p53 X3 + p52 X2 + p51 X1$
 - 4) $X6 = p65 X5 + p64 X4 + p63 X3$
- (Ahora se multiplica cada ecuación por Xi siendo i = n - 1.)

De la ecuación (1) :

$$X3 = p23 X2$$

$$X3 X2 = p23 X2 X2$$

($X3X2 = R32$ Coef. correlación entre X3 y X2)

$$R32 = p23 \times 1 = 0,23$$

Figura II

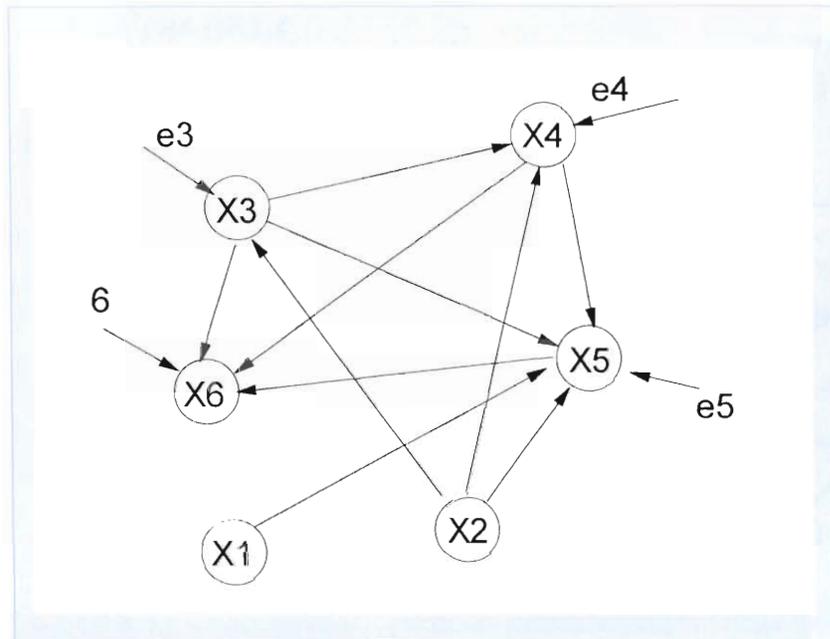
Tabla con los datos obtenidos de los partes de accidente:

Nº acciden	Mes	Edad	Dia	Antiguedad	Duración	Hora
	X1	X2	X3	X4	X5	X6
1	1	44	1	19	7	5
2	2	44	3	20	11	3
3	3	41	4	10	10	3
4	3	41	1	1	11	3
5	3	41	5	19	9	5
6	5	46	5	19	14	6
7	6	29	5	7	23	3
8	6	47	6	2	13	5
9	7	52	6	19	15	6
10	7	28	3	1	7	2
11	8	51	2	20	12	4
12	10	38	2	19	12	4
13	11	29	2	1	8	3
14	11	53	5	19	7	4
15	12	33	5	10	18	1
16	12	41	3	10	2	1
17	12	49	4	19	30	3
18	12	24	3	1	11	1

Figura III

	X6	X5	X4	X3	X2	X1	
	hora	duracion	antigued	dia	edad	mes	
X6	hora	1					
X5	duracion	.053	1				
X4	antigued	.525	.149	1			
X3	dia	.279	.361	.127	1		
X2	edad	.637	.069	.694	.232	1	
X1	mes	-.501	.167	-.159	.103	-.205	1

Figura IV



De la ecuación (2):

$$\begin{aligned} X4 &= p43 X3 + p24 X2 \\ X4 X3 &= p43 X3 X3 + p24 X2 X3 \\ X4 X2 &= p43 X3 X2 + p24 X2 X2 \\ R43 &= p43 + p24 R23 \\ R42 &= p43 R32 + p24 \quad (p43 = 0,032) \text{ y } (p24 = 0,697) \end{aligned}$$

De la ecuación (3):

$$\begin{aligned} X5 &= p54 X4 + p53 X3 + p52 X2 + p51 X1 \\ X5 X4 &= p54 X4 X4 + p53 X3 X4 + p52 X2 X4 + p51 X1 X4 \\ X5 X3 &= p54 X4 X3 + p53 X3 X3 + p52 X2 X3 + p51 X1 X3 \\ X5 X2 &= p54 X4 X2 + p53 X3 X2 + p52 X2 X2 + p51 X1 X2 \\ X5 X1 &= p54 X4 X1 + p53 X3 X1 + p52 X2 X1 + p51 X1 X1 \end{aligned}$$

Para calcular los coeficientes obtendremos mediante una hoja de cálculo la matriz inversa de la matriz 4x4 obteniendo otra matriz de dimensiones 4x4, multiplicando esta última matriz por el vector columna (el de los coeficientes de correlación) obtenemos otro vector columna cuyos elementos son los coeficientes de Wright buscados.

(Como las hojas de cálculo tipo Lotus no calculan determinantes sino matrices, tenemos que obtener primero la inversa y a continuación efectuar la multiplicación de la matriz obtenida por el vector columna) Los coeficientes buscados son:

$$\begin{aligned} p54 &= 0,22 \\ p53 &= 0,35 \\ p52 &= -0,13 \\ p51 &= 0,13 \end{aligned}$$

De la ecuación (4)

$$X6 = p65 X5 + p64 X4 + p63 X3$$

Multiplicando por X(n-1)

$$X6 X5 = p65 X5 X5 + p64 X4 X5 + p63 X3 X5$$

$$X6 X4 = p65 X5 X4 + p64 X4 X4 + p63 X3 X4$$

$$X6 X3 = p65 X5 X3 + p64 X4 X3 + p63 X3 X3$$

Sustituyendo:

$$\begin{pmatrix} 1 & 0,14 & 0,361 & p65 & 0,053 \\ 0,149 & 1 & 0,127 & p64 & 0,525 \\ 0,361 & 0,127 & 1 & p63 & 0,279 \end{pmatrix} \times \begin{pmatrix} p65 \\ p64 \\ p63 \end{pmatrix} = \begin{pmatrix} R54 \\ R53 \\ R52 \\ R51 \end{pmatrix}$$

Obteniendo la inversa de la matriz 3x3 y multiplicándola por el vector columna obtendremos otro vector columna cuyos elementos serán la solución de los coeficientes buscados.

$$\begin{aligned} (p65 &= -0,11), (p64 = 0,50) \text{ y} \\ (p63 &= 0,25) \end{aligned}$$

En forma matricial:

$$\begin{pmatrix} 1 & R34 & R24 & R14 \\ R43 & 1 & R23 & R13 \\ R42 & R32 & 1 & R12 \\ R41 & R31 & R21 & 1 \end{pmatrix} \times \begin{pmatrix} p54 \\ p53 \\ p52 \\ p51 \end{pmatrix} = \begin{pmatrix} R54 \\ R53 \\ R52 \\ R51 \end{pmatrix}$$

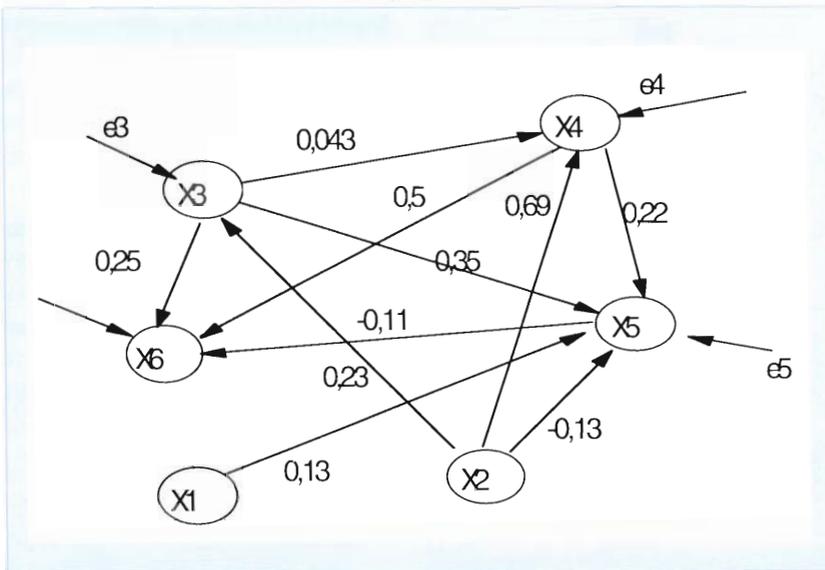
Sustituyendo valores

$$\begin{pmatrix} 1 & 0,127 & 0,694 & -0,159 \\ 0,127 & 1 & 0,232 & 0,103 \\ 0,694 & 0,232 & 1 & -0,205 \\ -0,159 & 0,103 & -0,205 & 1 \end{pmatrix} \times \begin{pmatrix} p54 \\ p53 \\ p52 \\ p51 \end{pmatrix} = \begin{pmatrix} 0,149 \\ 0,361 \\ 0,069 \\ 0,167 \end{pmatrix}$$

DOSSIER

A continuación se sustituyen los coeficientes obtenidos en el grafo:

Figura V



$$D_{ij} = \frac{\sum (X_{ij} - X_{jk})^2}{n}$$

Dij=distancia entre sujeto (i) y sujeto (j)

Xij=Valor de variable (i) del sujeto (i)

n = número de variables

A la vista del grafo y de los coeficientes se puede concluir que existe una dependencia fuerte entre la antigüedad y la edad, lo cual nos puede hacer pensar que la plantilla de la empresa está constituida por personal con contratos no-temporales. La edad del accidentado no parece que influya sobre la duración del accidente, en general y a partir de los datos con los que estamos trabajando y el modelo propuesto, se puede concluir que no existe una dependencia entre las variables estudiadas. Otra forma de estimar si existe alguna similitud entre los accidentes, en este caso de forma gráfica, consiste en la elaboración de un dendograma o representación gráfica de las distancias entre los sujetos accidentados, distancias que se calculan aplicando una métrica sobre los datos teniendo la precaución previa con objeto de hacerlos independientes de la escala de medida de dividirlos por la desviación típica. Las distancias entre los sujetos las obtendremos a partir de la siguiente relación:

Para cada una de las variables calcularemos su desviación típica y dividiremos cada dato por la desviación típica obteniendo la siguiente tabla:

Figura VI

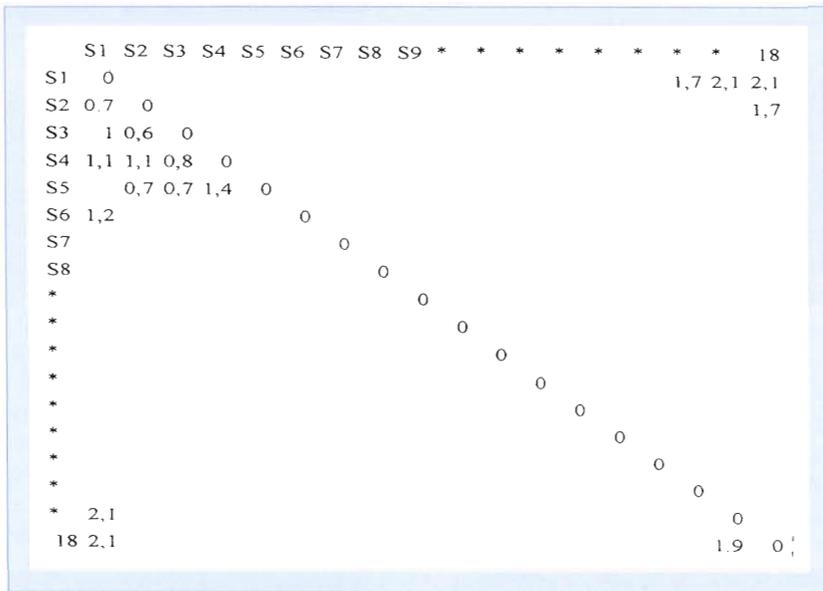
Sujeto	X1	X2	X3	X4	X5	X6
1	0,2	5	0,6	2,3	1	3,1
2	0,5	5	1,8	2,5	1,7	1,8
3	0,7	4,6	2,5	1,2	1,5	1,8
4	0,7	4,6	0,6	0,1	1,7	1,8
5	0,7	4,6	3,1	2,3	1,4	3,1
6	1,3	5,2	3,1	2,3	2,1	3,7
7	1,5	3,2	3,1	0,8	3,5	1,8
8	1,5	5,3	3,7	0,2	2	3,1
9	1,8	5,9	3,7	2,3	2,3	3,7
10	1,8	3,1	1,8	0,1	1	1,2
11	2,1	5,7	1,2	2,5	1,8	2,5
12	2,6	4,3	1,2	2,3	1,8	2,5
13	2,8	3,2	1,2	0,1	1,2	1,8
14	2,8	6	3,1	2,3	1	2,5
15	3,1	3,7	3,1	1,2	2,8	0,6
16	3,1	4,6	1,8	1,2	0,3	0,6
17	3,1	5,5	2,5	2,3	4,6	1,8
18	3,1	2,7	1,8	0,1	1,7	0,6

La distancia entre el sujeto (1) y el sujeto (2) será:

$$D12 = \text{SQR} \left(\frac{(0,2 - 0,5)^2 + (5-5)^2 + (0,6-1,8)^2 + (2,3-2,5)^2 + (1-1,7)^2 + (3,1-1,8)^2}{6} \right) = 0,78$$

De esta forma calcularíamos las 324 distancias entre los 18 sujetos accidentados.

Figura VII



Con estas distancias se dibuja el dendograma, donde sujetos iguales estarán juntos, los similares formaran clusters y los grupos de sujetos mas parecidos estarán mas agrupados que los disimilares. La no formación de grupos o racimos implicará que no hay mucha relación entre los accidentes.

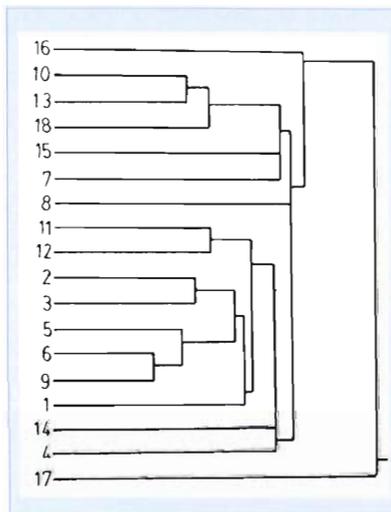
(Ver figura VIII)

Del dendograma podemos deducir que entre los accidentados o sujetos 6 y 9 hay mas similitud que entre el 14 y el 4 y que entre el 6, el 9 y el 5 hay menos diferencias que entre el 16 y el 8. En esta situación el dendograma no revela ningún grupo claro indicándonos que no existe una clara relación entre los accidentes; a esta conclusión ya habíamos llegado anteriormente mediante el análisis de senderos.

Por último, podríamos plantearnos la posibilidad de obtener una relación de tipo lineal (ecuación) que definiera o estimara el número de días de baja de un accidente (variable dependiente) en función de un conjunto de variables (variables independientes) obtenidas o extraídas del parte de accidente.

DENODOGRAMA ENTRE SUJETOS

Figura VIII



Con objeto de aumentar la varianza del modelo a buscar, vamos a «retocar» un poco la muestra (18 accidentes). En primer lugar y visto el dendograma, vamos a eliminar los accidentes o sujetos (16) y (17) que son respectivamente los

que presentan una duración menor y mayor en cuanto al número de días de baja.

Así hemos reducido la muestra a (16) sujetos.

Calculamos la media de duración de los días de baja de los restantes 16 sujetos y su desviación típica (media = 11,7 y desviación típica = 4,29) aplicamos el test de Grubbs, al sujeto (7) que tiene una duración de 23 días. (media - duración a testar) / Desv. típica = (11,7 - 23) / 4,2 = 2,67

La tabla de Grubbs da como valor límite para tomar el dato con n = 16 el valor de: 2,59 luego como el valor calculado supera al crítico (con una probabilidad del 95%) rechazamos o no consideramos al caso o sujeto número (7) y la muestra se nos queda con (15) sujetos.

Al aplicar un modelo de regresión lineal múltiple a las variables: X1, X2, X3, X4 y X5 tomando como variable dependiente la (X5) que representa el número de días de baja obtenemos la siguiente relación:

$$X5 = 0,344 X1 + 0,161 X2 + 0,656 X3 - 0,029 X4$$

(Donde X1, X2, X3 y X4 son respectivamente: mes, edad, día y antigüedad en el puesto)

El coeficiente de correlación múltiple es del orden de 0,968, el coeficiente de determinación tiene un valor de: 0,936 y el coeficiente de determinación corregido:

$$\text{Coef. deter. correg.} = R^2 - K(1 - R^2) / (N - K - 1) = 0,936 - 4(1 - 0,936) / (15 - 4 - 1) = 0,91$$

(Donde : R2 = R al cuadrado)

Número de grados de libertad = casos - variables correlacionadas = 15 - 5 = 10 con un valor crítico de 0,84 al 1%.

Como 0,968 > 0,84 se rechaza la Ho y se concluye que el coeficiente de correlación es significativo.

Así pues el modelo es capaz de explicar (ver coef. deter. correg.) el 91 % de la varianza, lo cual es bastante aceptable.

CONCLUSIÓN

- Las relaciones de dependencia entre las variables estudiadas es baja.
- La poca dependencia indica que no existe una casuística común.
- El dendograma no indica similitudes entre los accidentes.
- Existe una alta relación entre antigüedad y edad. El empleo es estable.