

Problemas de Análisis de la Varianza

1. El proceso de obtención de papel necesita grandes cantidades de agua. Una vez preparada la madera, ésta es normalmente calentada con agentes reductores y álcalis para disolver las fibras y obtener pulpa. Este proceso es muy contaminante. Se han tomado datos aleatorios a lo largo de un año de la cantidad de agua usada en el paso inicial de la obtención de pulpa por tonelada de madera en 4 trituradoras. Y se quiere ver si los resultados entre ellos varían:

A	B	C	D
22,0	18,6	23,1	20,3
18,8	16,0	18,0	18,5
19,5	17,7	20,0	19,6
23,7	14,6	21,0	22,8
21,9	19,9	23,7	23,8
19,9	18,6	22,4	21,4

Se supone que los datos proceden de una Normal.

2. Siguiendo el estudio de las características morfológicas de los árboles de Madrid, se quiere saber si en el riesgo de caída del árbol influye la distancia a la fachada.

Los datos son los siguientes:

Distancia	Riesgo
M	0,37
M	0,37
M	0,40
G	0,37
G	0,98
G	0,30

P: $d < 1\text{m}$; M: $1-2,5\text{ m}$; G: $d > 2,5\text{ m}$

Calcular la tabla de análisis de la varianza, y decidir si se acepta o no la influencia de la distancia en el riesgo. Para ello según los resultados de la tabla ANOVA escoged el valor del estadístico F y de la significancia p-valor, más adecuados entre los que se proponen:

1. $F=0,5$; $P\text{-valor}=0,52$
2. $F=0,62$; $p\text{-Valor}=0,47$
3. $F=1$; $p\text{-Valor}=0,009$

Problemas de Regresión

1. Se ha medido el contenido de oxígeno, Y, en miligramos/litro, de la laguna de Peñalara, en Madrid, a una profundidad de X metros, obteniéndose los siguientes datos:

Profundidad (m) X	Oxígeno(mgr/l) Y
15	5,5
20	5,3
30	5,2
40	5
50	4,3
60	1,4
70	0,1

- Calcular la recta de regresión.
- Estudiar la correlación entre variables.
- Para una profundidad entre 20 y 30 metros, ¿Qué contenido de Oxígeno se puede predecir?

2. Un botánico observa la longitud y la anchura de una muestra de 16 hojas de una determinada especie:

Hoja	Anchura (cm) X	Longitud (cm) Y
1	2,1	4,1
2	2,4	6
3	3,6	5,5
4	3,7	8,2
5	4,3	7,5
6	5,1	12,6
7	5,5	8,1
8	5,8	10,8
9	5,9	7,2
10	6,6	13,1
11	7,4	11,3
12	8,2	15,6
13	8,8	13,4
14	9	19
15	9,1	15,8
16	9,8	14,6

Suponiendo normalidad bivalente, se pide:

- ¿Son independientes la anchura y la longitud de estas hojas en la población? Tómese $\alpha=0,05$.
- ¿Puede admitirse que la correlación poblacional es $\rho=0,95$?

3.- Davison (“Update on Ozone Trends in California’s South Coast Air Basin”, Air and Water, 43, 226, 1993) estudió las concentraciones de ozono en la cuenca aérea de la costa sur de California, durante los años 1976 a 1991. Cree que la cantidad de días en que las concentraciones de ozono fueron mayores de 0.2 ppm (la respuesta) depende del índice meteorológico estacional, que es el promedio estacional de la temperatura con 850 milibares (el regresor). Los datos son los siguientes:

Año	Días	Índice
1976	91	16,7
1977	105	17,1
1978	106	18,2
1979	108	18,1
1970	88	17,2
1981	91	18,2
1982	58	16
1983	82	17,2
1984	81	18
1985	65	17,2
1986	61	16,9
1987	48	17,1
1988	61	18,2
1989	43	17,3
1990	33	17,5
1991	36	16,6

- Estimar la ecuación de regresión
- Probar la significancia de la regresión.
- Calcular los intervalos de confianza para la moda y la media y los de la predicción. Obtener los gráficos de estas bandas.
- ¿Cumplen los residuos las hipótesis del modelo lineal?

4.- La tabla siguiente muestra los datos obtenidos por Leffler y Nyholm (1999) de la concentración de cadmio en el riñón de los ratones de campo en función de la edad. Este mamífero ha disminuido su presencia en las Islas Británicas, y se piensa que es debido a la contaminación del agua. Se pide:

Edad en meses X	Cd ($\mu\text{gr}/\text{gr}$) Y
2,5	4,6
3	4
3,5	5,1
4	8,1
6	10,3

- Calcular la recta de regresión.
- ¿Se puede aceptar que el coeficiente de regresión sea 0 para un nivel de significación de 0.05?
- ¿Cuál es el valor estimado para una edad de 5 meses? Calcula los intervalos de confianza.