

APELLIDOS:

NOMBRE:

Advertencia: Indicar siempre con claridad de dónde proceden los valores que se utilicen en los diferentes cálculos.

Problema 1 (5 puntos)

Para comparar la calidad del agua de algunos ríos de la Comunidad de Madrid se puede utilizar el índice IBMWP, que se halla buscando diferentes especies de macroinvertebrados que nos encontramos a lo largo del río. Cada especie encontrada tiene un valor numérico, y con la suma total de dichos valores obtenemos el valor de dicho índice.

El estudio lo realizamos en tres ríos madrileños distintos: el Río Manzanares, el Río Mediano y el Río Jarama. El valor del índice puede variar en función del tramo del río en que busquemos los macroinvertebrados: en la cabecera del río, en el tramo medio o en la desembocadura. Los datos obtenidos son:

	R. Manzanares	R. Mediano	R. Jarama
Cabecera	14	13	13
Tramo medio	21	19	18
Desembocadura	10	9	7

Analizamos estos datos mediante SPSS y obtenemos los siguientes resultados:

Pruebas de los efectos inter-sujetos

Variable dependiente: Índice IBMWP

Origen	Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.
Modelo corregido	179,778 ^a	4	44,944	101,125	,000
Intersección	1708,444	1	1708,444	3844,000	,000
Río	8,222	2	4,111	9,250	,032
Tramo	171,556	2	85,778	193,000	,000
Error	1,778	4	,444		
Total	1890,000	9			
Total corregida	181,556	8			

a. R cuadrado = ,990 (R cuadrado corregida = ,980)

Comparaciones múltiples

Variable dependiente: Índice IBMWP
Bonferroni

(I)Tramo del río	(J)Tramo del río	Diferencia de medias (I-J)	Error típ.	Sig.	Intervalo de confianza 99%	
					Límite inferior	Límite superior
Cabecera	Tramo medio	-6,00*	,544	,001	-9,40	-2,60
	Desembocadura	4,67*	,544	,003	1,26	8,07
Tramo medio	Cabecera	6,00*	,544	,001	2,60	9,40
	Desembocadura	10,67*	,544	,000	7,26	14,07
Desembocadura	Cabecera	-4,67*	,544	,003	-8,07	-1,26
	Tramo medio	-10,67*	,544	,000	-14,07	-7,26

Basadas en las medias observadas.

El término de error es la media cuadrática(Error) = ,444.

*. La diferencia de medias es significativa al nivel ,01.

Comparaciones múltiples

Variable dependiente: Índice IBMWP
Bonferroni

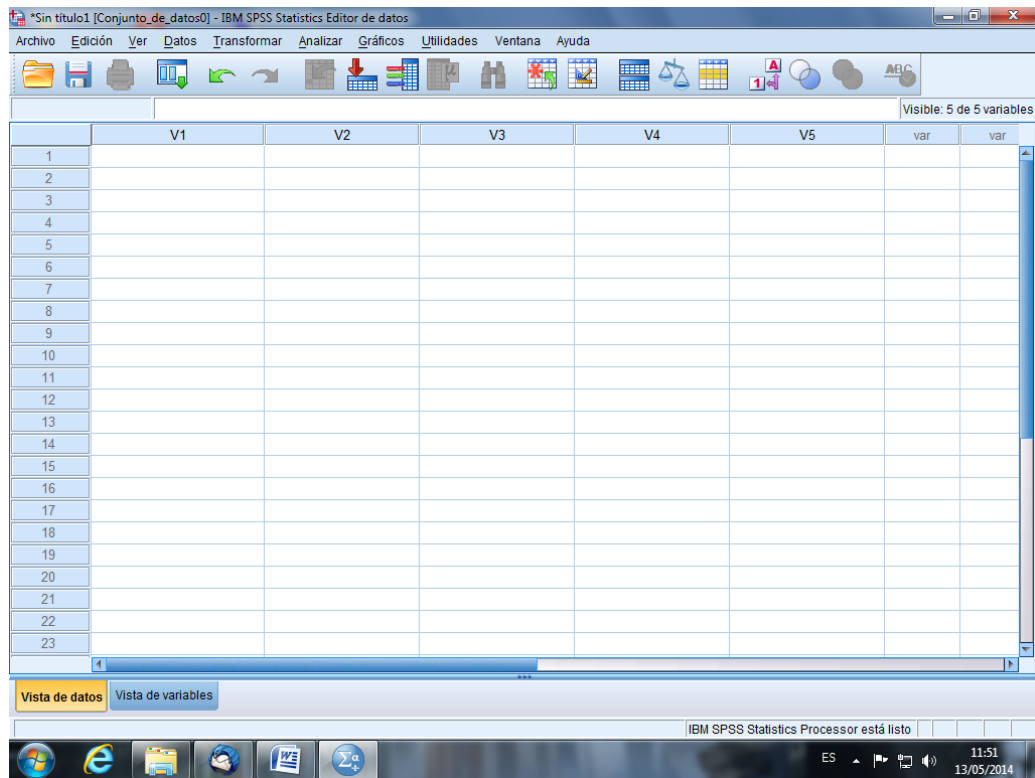
(I)Río	(J)Río	Diferencia de medias (I-J)	Error típ.	Sig.	Intervalo de confianza 99%	
					Límite inferior	Límite superior
Río Manzanares	Río Mediano	1,33	,544	,211	-2,07	4,74
	Río Jarama	2,33	,544	,038	-1,07	5,74
Río Mediano	Río Manzanares	-1,33	,544	,211	-4,74	2,07
	Río Jarama	1,00	,544	,420	-2,40	4,40
Río Jarama	Río Manzanares	-2,33	,544	,038	-5,74	1,07
	Río Mediano	-1,00	,544	,420	-4,40	2,40

Basadas en las medias observadas.

El término de error es la media cuadrática(Error) = ,444.

- (a) Describir con detalle el modelo estadístico empleado, y estimar el efecto adicional que aporta la desembocadura sobre el índice medio de calidad (indicando a qué parámetro corresponde).

- (b) Indicar claramente en la siguiente captura de pantalla, cómo se han introducido los datos para poder aplicar el modelo utilizado:



- (c) ¿Influye el tramo sobre el índice de calidad del agua? ¿Influye el río? Plantear los contrastes, indicando claramente la hipótesis nula y la hipótesis alternativa, y responder a un nivel de significación del 1%.

(d) Si comparamos los diferentes tramos de dos en dos (con un 1% de nivel de significación conjunto), ¿qué conclusiones obtenemos?

(e) Calcular un intervalo (al 90% de confianza) para estimar la diferencia entre los índices medios de calidad del Manzanares y del Jarama.

Problema 2 (5 puntos)

Antes de lanzar al mercado un producto comercial dirigido a la reducción del colesterol, se quiere llevar a cabo un estudio estadístico. Un grupo de 9 personas consume, durante 3 semanas, diferentes cantidades semanales del producto y, al cabo de las 3 semanas se mide la reducción de colesterol que se ha conseguido. Llamaremos Y a la reducción de colesterol conseguida (en mg/dl), y llamaremos X al número total de envases monodosis consumidos por cada persona en esas 3 semanas. Los datos obtenidos se analizaron con SPSS mediante un modelo de regresión lineal, obteniéndose los siguientes resultados:

ANOVA^a

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	209,067	1	209,067	24,833	,002 ^b
	Residual	58,933	7	8,419		
	Total	268,000	8			

a. Variable dependiente: Reducción de colesterol (en mg/dl)

b. Variables predictoras: (Constante), Número total de envases

Coefficientes^a

Modelo		Coefficients no estandarizados		Coefficients tipificados	t	Sig.
		B	Error típ.	Beta		
1	(Constante)	3,267	3,688		,886	,405
	Número total de envases	,933	,187	,883	4,983	,002

a. Variable dependiente: Reducción de colesterol (en mg/dl)

(a) Escribir el modelo lineal ajustado que hemos obtenido a partir de esos datos. Calcular e interpretar el coeficiente de determinación.

(b) Obtener el intervalo de confianza para estimar (al 95%) el coeficiente que multiplica a la variable independiente .

- (c) El número total de envases consumidos, ¿tiene una influencia significativa sobre Y? Plantear claramente el contraste de hipótesis, y responder a un nivel de significación del 5%.

A continuación, se muestran los resultados sobre tres modelos de regresión:

$$Y = \text{cte.} + b_1 X \quad ; \quad Y = \text{cte.} + b_1 \ln X \quad (\ln = \text{logaritmo neperiano}) \quad ; \quad Y = \text{cte.} \exp(b_1 X)$$

Resumen del modelo y estimaciones de los parámetros

Variable dependiente: Reducción de colesterol (en mg/dl)

Ecuación	Resumen del modelo					Estimaciones de los parámetros	
	R cuadrado	F	gl1	gl2	Sig.	Constante	b1
Lineal	,780	24,833	1	7	,002	3,267	,933
Logarítmica	,849	39,248	1	7	,000	-29,851	17,506
Exponencial	,731	18,983	1	7	,003	7,524	,052

La variable independiente es Número total de envases.

- (d) De los tres modelos de regresión considerados, ¿cuál parece mejor y por qué?

- (e) Estimar la reducción media de colesterol obtenida con el modelo de regresión logarítmica y con el modelo de regresión exponencial, cuando se consume un total de 20 envases monodosis durante las tres semanas.