



Asignatura..... **Análisis de datos**..... Grupo.....  
 Apellidos..... Nombre.....  
 Ejercicio del día ..... **24 de junio de 2015**.....

1.— Una agencia estatal recoge todos los años datos sobre los contenidos en nicotina y alquitrán (en mg), el peso (en g) y la emisión total de CO<sub>2</sub> (en mg) de las marcas comerciales de cigarrillos. Analizados los datos de un determinado año, en el que se estudiaron 25 presentaciones comerciales se obtuvieron los datos siguientes.

**Estadísticos descriptivos**

	Media	Desviación típica	N
alquitrán	12,216	5,6658	25
nicotina	,8764	,35406	25
peso	,970284	,0877215	25
CO2	12,528	4,7397	25

**Correlaciones**

		alquitrán	nicotina	peso	CO2
alquitrán	Correlación de Pearson	1	,977**	,491*	,957**
	Sig. (bilateral)		,000	,013	,000
	N	25	25	25	25
nicotina	Correlación de Pearson	,977**	1	,500*	,926**
	Sig. (bilateral)	,000		,011	,000
	N	25	25	25	25
peso	Correlación de Pearson	,491*	,500*	1	,464*
	Sig. (bilateral)	,013	,011		,019
	N	25	25	25	25
CO2	Correlación de Pearson	,957**	,926**	,464*	1
	Sig. (bilateral)	,000	,000	,019	
	N	25	25	25	25

\*\* . La correlación es significativa al nivel 0,01 (bilateral).

\* . La correlación es significativa al nivel 0,05 (bilateral).

Si se tratan de explicar las emisiones de CO<sub>2</sub> por medio de las variables «alquitrán», «nicotina» y «peso» se obtienen los resultados

**Resumen del modelo<sup>b</sup>**

Modelo	R	R cuadrado	R cuadrado corregido	Error típ. de la estimación
1	,958 <sup>a</sup>	,919	,907	1,4457

a. Variables predictoras: (Constante), alquitrán, peso, nicotina

b. Variable dependiente: CO2

**ANOVA<sup>a</sup>**

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	495,258	3	165,086	78,984	,000 <sup>b</sup>
	Residual	43,893	21	2,090		
	Total	539,150	24			

a. Variable dependiente: CO2

b. Variables predictoras: (Constante), alquitrán, peso, nicotina

**Coefficientes<sup>a</sup>**

Modelo		Coefficients no estandarizados		Coefficients tipificados	t	Sig.
		B	Error típ.	Beta		
1	(Constante)	3,202	3,462		,925	,365
	peso	-,130	3,885	-,002	-,034	,974
	nicotina	-2,632	3,901	-,197	-,675	,507
	alquitrán	,963	,242	1,151	3,974	,001

a. Variable dependiente: CO2

a) Describe el modelo utilizado y sus requisitos previos ¿Qué se puede decir de la colinealidad?

b) Escribe la hipótesis nula y alternativa sobre el contraste del coeficiente de la variable «nicotina». ¿Qué decisión se toma en este contraste si se establece una significación de 0,05?

c) Calcula el intervalo de confianza del 90 % para el coeficiente de la variable «peso». ¿Puedo afirmar que este coeficiente es significativamente negativo?

Dados los problemas observados, se decide utilizar como variables explicativas el contenido de alquitrán, el peso del cigarrillo y el producto de ambas variables (que aparece en las siguientes tablas como «AxP»). Se obtiene los siguientes resultados.

Resumen del modelo<sup>b</sup>

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,974 <sup>a</sup>	,948	,940	1,1577

a. Variables predictoras: (Constante), AxP, peso, alquitran

b. Variable dependiente: CO2

Coefficientes<sup>a</sup>

Modelo		Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
		B	Error típ.	Beta		
1	(Constante)	-10,320	4,707		-2,192	,040
	alquitran	1,884	,310	2,252	6,085	,000
	peso	13,220	4,949	,245	2,671	,014
	AxP	-1,075	,305	-1,442	-3,530	,002

a. Variable dependiente: CO2

d) ¿Qué conclusiones se obtienen ahora? Justifica tu respuesta.

e) Estima, utilizando por separado ambos modelos, la cantidad de CO<sub>2</sub> emitida por un cigarrillo de peso 1 g, con un contenido de alquitrán de 10 mg, y —solamente para el primer modelo— un contenido de nicotina de 0,75 mg. Indica cual es, en cada uno de los dos modelos, el incremento en emisión de CO<sub>2</sub> cuando, para un peso fijo del cigarrillo, el contenido en alquitrán se incrementa en 1 mg.

Primer modelo:

Segundo modelo:

2.— Impacto de la vitamina B en el tamaño del riñón. Un lote de 28 ratas se clasificó en dos grupos de 14 ratas cada uno según su tamaño (Delgada, Obesa).

Cada uno de los dos grupos de 14 ratas se dividió aleatoriamente en dos grupos de 7 ratas que se alimentaron durante 20 días con dietas diferentes (Normal, Vitamina B).

Al final del tratamiento se pesó el hígado de cada una de las ratas. Los resultados obtenidos se trataron con SPSS y se obtuvieron los siguientes resultados.

**Estadísticos descriptivos**

Variable dependiente: PESO

Tamaño	Dieta	Media	Desviación típica	N
Delgada	Normal	1,6414	,16658	7
	Vitamina B	1,5271	,12148	7
	Total	1,5843	,15210	14
Obesa	Normal	2,6429	,34340	7
	Vitamina B	2,6729	,26107	7
	Total	2,6579	,29347	14
Total	Normal	2,1421	,58072	14
	Vitamina B	2,1000	,62584	14
	Total	2,1211	,59280	28

**Pruebas de los efectos inter-sujetos**

Variable dependiente: PESO

Origen	Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.
Modelo corregido	8,117 <sup>a</sup>	3	2,706	47,345	,000
Intersección	125,970	1	125,970	2204,345	,000
TAMANNO	8,068	1	8,068	141,179	,000
DIETA	,012	1	,012	,218	,645
TAMANNO * DIETA	,036	1	,036	,638	,432
Error	1,372	24	,057		
Total	135,459	28			
Total corregida	9,488	27			

a. R cuadrado = ,855 (R cuadrado corregida = ,837)

a) ¿Qué modelo se ha aplicado para obtener la tabla ANOVA anterior? Describe todos sus elementos y relaciones.

b) En función de los parámetros del modelo descrito en a), escribe las hipótesis nula y alternativa del contraste sobre la interacción de los factores.

Al nivel de significación  $\alpha = 0,1$ , decide este contraste. Indica que valor o valores usas para tomar esta decisión.

c) Da un intervalo de confianza del 95% para la diferencia de las medias de peso del hígado entre ratas Obesas y ratas Delgadas.

Un análisis de residuos ha dado los siguientes resultados:

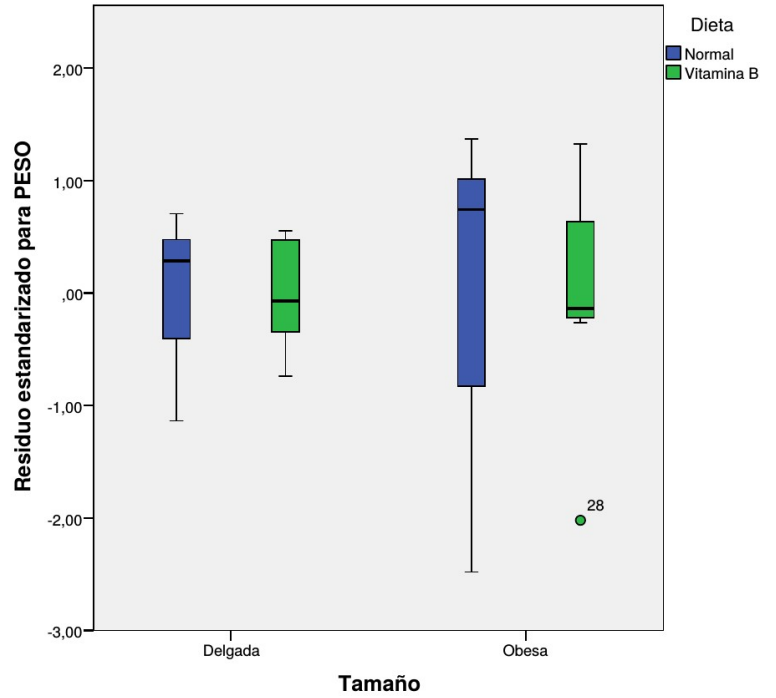
**Contraste de Levene sobre la igualdad de las varianzas error<sup>a</sup>**

Variable dependiente: PESO

F	gl1	gl2	Sig.
2,632	3	24	,073

Contrasta la hipótesis nula de que la varianza error de la variable dependiente es igual a lo largo de todos los grupos.

a. Diseño: Intersección + TAMANNO + DIETA + TAMANNO \* DIETA



d) Razona si los resultados del contraste de Levene y lo observado en los diagramas de cajas son coherentes. ¿Qué conclusión se alcanza?

e) Construye la tabla ANOVA que se obtendría con los mismos datos si se ignora el factor DIETA (escribe la  $F$  crítica al nivel  $\alpha = 0,01$ ). ¿Qué conclusión se obtiene?

Fuente	Suma de cuadrados	g.l.	Media cuadrática	$F$	$F$ -critica
TAMANNO					
Error					
Total					