

CUESTIÓN 1: Dos muestras independientes, ambas de tamaño 7, de sendas distribuciones poblacionales normales de varianza común σ^2 arrojan medias muestrales de 4.8 y 5.4, y varianzas muestrales de 8.38 y 7.62, respectivamente. Encontrar la expresión que permitiría obtener un intervalo de confianza para $\mu_1 - \mu_2$ (diferencia de sus medias poblacionales) al nivel 0.95.

CUESTIÓN 2: Se supone que el número de erratas por página de un libro sigue una distribución de Poisson. Elegidas al azar 95 páginas se obtuvo que había 0, 1, 2, 3, 4, 5 en 40, 30, 15, 7, 1 y 0 páginas respectivamente. ¿Contiene la muestra evidencias estadísticamente significativas para rechazar dicho supuesto? Justifica la respuesta.

CUESTIÓN 3: Sea una población de 3.245 empresas dedicadas al sector de comercio estratificadas en 3 grupos de tamaño, según el número de asalariados. Los datos disponibles son los siguientes:

Estrato	N_h	X_h	S_h
1 (entre 0 y 9 asalariados)	3.200	3.400	1,5
2 (entre 10 y 49 asalariados)	40	650	10,2
1 (de 50 o más asalariados)	5	580	150

Donde:

N_h = Número de empresas en el estrato h

X_h = Total de asalariados en el estrato h

S_h = Cuasivarianza poblacional de X en el estrato h

El tamaño muestral es de 50 empresas y el muestreo es aleatorio simple sin reposición. Se pide:

- Calcular el tamaño muestral por estrato, usando la afijación de mínima varianza o de Neyman.
- Calcular la varianza y el coeficiente de variación para el estimador insesgado del total de asalariados.

CUESTIÓN 4: Se quiere conocer la superficie dedicada a la plantación de pinos de una región. Dicha región, con un total de 50.000 Km^2 , se divide en 100 áreas o conglomerados. Se extrae una muestra de 5 conglomerados con reemplazamiento y con probabilidades proporcionales a sus superficies. Las proporciones de superficie dedicadas a la plantación de pinos, en cada uno de los conglomerados de la muestra, son las siguientes:

Conglomerado	1	2	3	4	5
Proporción	0,05	0,2	0,1	0,15	0,25

Se pide calcular un estimador insesgado de la superficie total dedicada a la plantación de pinos y su error de muestreo relativo o coeficiente de variación.

CUESTIÓN 5: En una determinada economía, para la rama de actividad A, se tiene la siguiente información:

	año t	año t+1
Producción (p.b.) a precios corrientes (millones de euros)	450.000	482.000
Consumos intermedios (p.a.) a precios corrientes (millones de euros)	275.000	292.000
Índice de precios de la producción de la rama (en %)	100	105
Índice de precios de los consumos intermedios de la rama (en %)	100	103

Por otra parte, el PIB (p.m.) a precios corrientes de esa economía en el año t fue de 1.118.000 millones de euros; la tasa de variación anual de este agregado a precios corrientes entre t y t+1 fue del 4,0% mientras que su tasa de variación anual en volumen fue del 3,1%. Se pide:

a) Calcule la tasa de variación del valor añadido (p.b) a precios corrientes de la rama de actividad A entre t y t+1. Obtenga asimismo la tasa de variación en volumen del valor añadido (p.b) entre t y t+1.

b) Calcule el valor del PIB (p.m) en t+1 a precios corrientes y la variación anual entre t y t+1 del deflactor implícito del PIB (p.m)

Nota: p.b.(precio básico), p.a (precio de adquisición), p.m.(precios de mercado)

CUESTIÓN 6: Se dispone de la siguiente información de los principales agregados de oferta y rentas de una economía (en millones de euros):

	Cod. SEC 2010	
Producción de bienes y servicios (a precios básicos)	P1	2.000.000
Consumo intermedio (a precios de adquisición)	P2	1.020.000
Impuestos sobre la producción y las importaciones	D.2	112.000
Impuestos sobre los productos	D.21	90.000
Subvenciones	D.3	18.400
Subvenciones sobre los productos	D.31	6.000
Excedente de explotación de renta mixta	B.2.G+B.3.G	445.000

Por otra parte, de la cuenta del resto del mundo se obtienen los datos siguientes (en millones de euros):

	Cod. SEC 2010	
Remuneración de asalariados recibida del resto del mundo	D.1	1.100
Remuneración de asalariados pagada al resto del mundo	D.1	300
Impuestos sobre la producción e importaciones pagados al resto del mundo	D.2	1.500
Subvenciones recibidas del resto del mundo	D.3	6.000
Otras transferencias corrientes recibidas del resto del mundo	D.7	2.000
Rentas de la propiedad recibidas del resto del mundo	D.4	45.000
Rentas de la propiedad pagadas al resto del mundo	D.4	66.000
Otras transferencias corrientes pagadas al resto del mundo	D.7	1.500

Calcule, a partir de esta información, el valor añadido total a precios básicos, el PIB a precios de mercado, la remuneración de los asalariados y la renta nacional bruta de la economía.

CUESTIÓN 7: En un Estado (E) de nuestro interés acontece que los accidentes de tráfico son la causa principal de muerte para jóvenes entre 5 y 32 años. A través de varias políticas de gasto, el gobierno central hace un tiempo promulgó una ley de uso obligatorio del cinturón de seguridad, y así reducir el número de muertes y de lesiones graves. Su departamento de análisis econométrico ubicado en su Instituto Nacional (INE) está interesado en examinar cómo de efectiva ha sido esta ley. Para ello, tiene disponible un conjunto de datos de panel de 50 unidades político-administrativas del país (piénsese en provincias, por ejemplo) para los años 1983-1997. Su conjunto de datos incluye las siguientes variables:

- *fatalidades* es el número de muertes por miles de millas de tráfico
- *CS_uso* es el nivel de uso del cinturón de seguridad
- *velo90* es una dummy = 1 si límite de velocidad es de 90 kms por hora, = 0 al contrario
- *velo120* es una dummy = 1 si límite de velocidad es de 120 kms por hora, = 0 al contrario
- *ba08* es una dummy = 1 si el límite de alcohol en la sangre es $\leq 0,08\%$, = 0 al contrario
- *beber – edad18* es una dummy = 1 si tiene 18 años de edad para beber, = 0 al contrario
- *ing* es el ingreso per cápita
- *edad* media de edad
- *provincia* es un conjunto de dummies de provincias
- *año* es un conjunto de dummies de años

La siguiente tabla contiene los resultados de varias regresiones (MCO agrupados, MCO con efectos fijos de estado, MC generalizados con efectos aleatorios de estado y MCO con efectos fijos de estado y año). Las siguientes preguntas se basan en estos resultados.

	1	2	3	4
<i>CS_uso</i>	4,07 (1,22)	-5,77 (1,215)	-4,50 (1,12)	-3,72 (1,13)
<i>velo90</i>	0,148 (0,403)	-0,425 (0,334)	-0,341 (0,337)	-0,783 (0,424)
<i>velo120</i>	2,40 (0,511)	1,23 (0,329)	1,34 (0,328)	0,804 (0,340)
<i>ba08</i>	-1,92 (0,445)	-1,38 (0,373)	-1,36 (0,367)	-0,822 (0,352)
<i>edad – beber18</i>	0,079 (0,876)	0,745 (0,507)	0,767 (0,510)	-1,13 (0,535)
<i>ln(ing)</i>	-18,1 (0,931)	-13,5 (1,42)	-12,6 (1,14)	6,26 (3,86)
<i>edad</i>	-0,007 (0,109)	0,979 (0,382)	0,232 (0,239)	1,32 (0,383)
<i>constante</i>	196,5 (8,22)		137,9 (8,92)	
Efectos provincia	No	FE	RE	FE
Efectos años	No	No	No	FE
R^2	0.544	0.874	0.683	0.897

FE (Efectos fijos); RE(Efectos aleatorios)

1. Centrándonos en los resultados de la regresión MCO de datos agrupados (fusionados) en la columna 1,

- a) ¿la regresión estimada sugiere que un mayor uso del cinturón de seguridad reduce significativamente las muertes?
- b) ¿indique si este resultado tiene sentido? Si es así, explica por qué. Si no, explica que piensas que está pasando aquí.
2. Las columnas 2 y 3 contemplan las estimaciones por efectos fijos de provincia efectos aleatorios de provincia.
- a) ¿Qué conclusión puede sacarse con respecto al impacto de nivel de uso del cinturón de seguridad en las muertes cuando agregamos los efectos fijos?
3. Los resultados de la prueba de Hausman para los modelos de las columnas 3 y 4 son los siguientes

	Coeficientes			
	(b) fijos	(B)	(b-B) diferencia	$(\text{diag}(\text{var}_-(b)-\text{var}_-(B)))^{-1}$ S.E.
CB_uso	-5.774782	-4.503969	-1.270813	0.2698476
velo90	-0.4250387	-0.3405939	-0.0844448	0.0644951
velo120	1.23329	1.335134	-0.1018444	0.0209829
ba08	-1.377456	-1.364296	-0.0131597	0.0648635
edad-beber18	0.7453195	0.766994	-0.0216745	
ln(ing)	-13.5144	-12.61544	-0.8989608	0.8380231
edad	0.9786802	0.2318357	0.7468445	0.2976956

b=consistente bajo H_0 y H_a ;

B=inconsistente bajo H_a , eficiente bajo H_0 ;

Test: H_0 : discrepancias de coeficiente no sistemáticas

$$\begin{aligned}
 \text{chi2}(7) &= (b - B)'[(\text{var}(b) - \text{var}(B))^{-1}](b - B) \\
 &= 26,62 \\
 \text{Prob} > \text{chi2} &= 0,0004
 \end{aligned}$$

Indica qué conclusión principal puedes sacar respecto de la información suministrada de cara a plantear el modelo más adecuado.

4. El modelo en la columna 4 agrega efectos fijos anuales al modelo de la columna 2. Una prueba F de la significación conjunta de estas dummies de 14 años nos da el estadístico F de 8.85. ¿Qué concluye acerca de la significación conjunta de las dummies del año? ¿Los resultados con respecto al impacto del uso del cinturón de seguridad cambian ahora que hemos agregado EF de los años?

CUESTIÓN 8: Está interesado en estimar una demanda de pescado vendido en el Mercado de Luarca, para ello usted va al mercado y recoge el precio diario y la cantidad vendida de 97 días consecutivos (dado que el mercado está cerrado los fines de semana, recopila datos de lunes a viernes). Específicamente, tiene datos de las siguientes variables:

- *totqty* - la cantidad total de pescado vendido ese día
- *avgprc* - el precio promedio del pescado vendido ese día
- *lun* - una dummy =1 por si el día es lunes
- *mar* - una dummy =1 por si el día es martes
- *mie* - una dummy =1 por si el día es miércoles
- *jue* - una dummy =1 por si el día es jueves
- *wave2* - la altura máxima promedio de olas durante dos días anteriores a los datos de precio y cantidad
- *wave3* - la altura media promedio de ola durante tres y cuatro días anteriores a los datos de precio y cantidad

Nota: aunque usemos subíndices de tiempo a lo largo de esta pregunta, no utilizaremos métodos de series temporales. También mantendremos la suposición de que todos los errores son homocédasticos.

1. Suponga que la ecuación de demanda se puede escribir para cada período de tiempo como $\ln(\text{totqty}_t) = \beta_0 + \beta_1 \ln(\text{avgprc}_t) + \beta_2 \text{lun}_t + \beta_3 \text{mar}_t + \beta_4 \text{mie}_t + \beta_5 \text{jue}_t + u_t$ la demanda puede variar durante días de la semana. ¿Por qué no es apropiado usar MCO para estimar esta ecuación de demanda? ¿Qué información adicional necesitamos para tener estimadores consistentes de los parámetros de ecuación de demanda?

2. Las variables wave2_t y wave3_t son medidas de las alturas de las olas oceánicas en los últimos días. ¿Qué dos suposiciones necesitamos para usar wave2_t y wave3_t como instrumentos para $\ln(\text{avgprc}_t)$ al estimar la ecuación de la demanda? Asegúrese de analizar cómo estas suposiciones están relacionadas con las ecuaciones de demanda y oferta.

3. La primera etapa de una regresión de mínimos cuadrados de dos etapas nos da los siguientes resultados:

$$\ln(\text{avgprc}) = -1,02 - 0,012\text{lun}_t - 0,0090\text{mar}_t + 0,051\text{mie}_t + 0,124\text{jue}_t + 0,094\text{wave2}_t + 0,053\text{wave3}_t$$

(0,14)
(0,114)
(0,1119)
(0,112)
(0,111)
(0,021)
(0,020)

$$\hat{R}^2 = 0,165$$

y la prueba de la significación conjunta de wave2_t y wave3_t nos da un estadístico F: $F - \text{stat} = 19,1$, mientras la prueba de significación conjunta de las dummies de día de la semana da $F - \text{stat} = 0,53$. ¿Son wave2_t y wave3_t individualmente significativos al nivel de 1%? ¿Qué revelan los resultados de la regresión de esta primera etapa sobre nuestros instrumentos?

4. Estimación de la ecuación de demanda por mínimos cuadrados en dos etapas nos da los siguientes resultados:

$$\ln(\text{totqty}_t) = 8,16 - 0,816\ln(\text{avgprc}_t) - 0,307\text{lun}_t - 0,685\text{mar}_t + 0,521\text{mie}_t + 0,095\text{jue}_t$$

(0,18)
(0,327)
(0,229)
(0,226)
(0,224)
(0,225)

donde los errores estándar entre paréntesis son los correctos (es decir, toman en cuenta el procedimiento de dos etapas). ¿Cuál es la interpretación del coeficiente en $\ln(\text{avgprc}_t)$?

¿Su magnitud parece razonable? Construya un intervalo de confianza del 95 % para este coeficiente.

5. Dado que tenemos dos instrumentos y una variable endógena, la ecuación de demanda está sobre identificada. La prueba de sobre identificación de restricciones nos da un estadístico F de 0,013. ¿Qué concluye?

6. Dado que la ecuación de la oferta (no especificada) evidentemente depende de las variables de olas(wave), ¿qué dos suposiciones deberíamos hacer para estimar la elasticidad del precio de la oferta?

7. Aquí están los resultados de la estimación de una posible ecuación de oferta para esta industria por mínimos cuadrados en dos etapas:

$$\ln(\text{totqty}_t) = 10,82 + 2,13\ln(\text{avgprc}_t) - 0,267\text{wave2}_t - 0,169\text{wave3}_t$$

(2,23) (2,24) (0,212) (0,139)

¿Cuál es la interpretación del coeficiente en $\ln(\text{avgprc}_t)$? ¿Es estadísticamente significativo al nivel de 10 %?

CUESTIÓN 9: Dada la siguiente tabla de mortalidad:

	Tasa de mortalidad	Promedio de años vividos el último año de vida	Riesgo de muerte	Supervivientes	Defunciones teóricas	Población estacionaria	Tiempo por vivir	Esperanza de vida
98 años	296,19	0,4758	256,38	6.263,77	1.605,93	5.421,93	16.422,50	2,622
99 años		0,4688			1.223,66			
100 y más años		2,0362						

Se pide:

- Calcular los supervivientes a las edades de 99 años y de 100 y más años.
- Calcular el riesgo de muerte a las edades de 99 años y de 100 y más años.
- Calcular la población estacionaria a las edades de 99 años y de 100 y más años.
- Calcular el tiempo por vivir a las edades de 99 años y de 100 y más años.
- Calcular la esperanza de vida a las edades de 99 años y de 100 y más años.
- Calcular la tasa de mortalidad a las edades de 99 años y de 100 y más años.

CUESTIÓN 10: Sea un país con dos regiones, A y B, y considere la siguiente matriz migratoria por regiones en el periodo 2001-2011:

Región de residencia 2001	Región de residencia 2011		
	Región A	Región B	Total
Región A	1.244,4	14,4	1.258,8
Región B	18,2	889,4	907,6
En el extranjero	72,2	99,3	171,5
No aplicable	130,1	123,2	253,3
Total	1.464,9	1.126,3	2.591,2

Se pide:

- El número total de emigrantes durante el periodo, la proporción de emigración y la tasa de emigración en las regiones A y B.
- El número total de inmigrantes durante el periodo, la proporción de inmigración y el índice de atracción en las regiones A y B.
- El saldo migratorio y las tasas de migración brutas y netas durante el periodo en las regiones A y B.
- El índice de efectividad migratoria y el índice de migración diferencial en las regiones A y B.

Nota: 'No aplicable' recoge la población nacida en el período intercensal 2001 - 2011.

1. Enumerar las ventanas que existen en una sesión de SAS, indicando brevemente para qué sirve cada una de ellas.
2. Describir brevemente la estructura general de un fichero SAS.
3. ¿Qué es una librería SAS? Escribese un ejemplo de sentencia de creación de una librería SAS.
4. Enumerar los componentes principales de un programa SAS, indicando brevemente qué función cumple cada componente.
5. Proporcionar tres ejemplos de operadores entre variables alfanuméricas en SAS.
6. Escribese un paso **DATA** para leer un fichero de texto (sin formato) de nombre `file.txt` con variables `var1` y `var2` alfanuméricas en columnas separadas por una coma (,). El fichero `file.txt` se encuentra en la ruta `C:\Documents`.
7. Un fichero SAS de nombre `inputFile` contiene dos variables numéricas `var1` y `var2`. ¿Escribese un breve código SAS para generar un nuevo fichero SAS de nombre `outputFile` que contenga la nueva variable `sumVar` que resulta de la suma de `var1` y `var2` para cada registro.
8. Explicar brevemente la diferencia entre estos dos pasos **DATA**:

<code>DATA outputFile;</code>	<code>DATA outputFile (DROP = var);</code>
<code>SET inputFile (DROP = var);</code>	<code>SET inputFile;</code>
<code>RUN;</code>	<code>RUN;</code>
9. Dados dos ficheros SAS `file1` y `file2` ordenados por la variable `ID`, escribir un paso **DATA** para obtener un fichero SAS de nombre `output` que solo contenga los registros del fichero `file1` que no estén en el fichero `file2`.
10. ¿Qué sentencias pueden especificarse en un procedimiento (**PROC PRINT**)?

