

Examen Final de Estadística I, 26 de mayo de 2010.
Grados ADE, ADE-DER, ADE-INF, ECO, ECO-DER.

- NORMAS:** 1) Entregar cada problema en un cuadernillo distinto, aunque esté en blanco.
2) Realizar los cálculos con al menos dos cifras decimales significativas.
3) No se podrá abandonar el examen hasta transcurridos 30 minutos después de haber empezado.
4) No está permitido salir del aula sin entregar el examen, aunque esté en blanco.

Problema 1 (10 puntos) Se realiza un experimento mediante el cual personas que han sido condenadas por pequeñas faltas deben realizar pagos mensuales a las víctimas según el criterio de un juez. La siguiente tabla contiene el salario mensual (x) de 10 condenados y los pagos (en euros) realizados a las víctimas (y):

<i>Salarios</i>	300	880	1000	1540	1560	1600	1600	2200	3200	6000
<i>Pagos</i>	200	380	400	200	800	600	800	1000	1600	2700

Se dispone además de la siguiente información: $\bar{x} = 1988$, $\bar{y} = 868$, $s_x^2 = 2594240$, $s_y^2 = 593351.1$ y $s_{xy} = 1206062$.

- (2 puntos) Representa el diagrama de dispersión (manteniendo la escala) que muestre los pagos como función de los salarios. Calcula e interpreta el coeficiente de correlación lineal de Pearson. Razona si crees que, aparentemente, existe relación lineal entre ambas variables.
- (2 puntos) Calcula la recta de regresión de y sobre x . Determina cuál es la pendiente de la recta e interpreta su valor. Representa la recta de regresión sobre el gráfico del apartado a) explicando detalladamente cómo hacer esta representación.
- (2 puntos) Calcula el coeficiente de determinación y el porcentaje de variabilidad no explicado por el modelo de la recta de regresión. Justifica si consideras adecuado este modelo.
- (2 puntos) Si el salario de una persona culpable de cierto delito es de 1400 euros, ¿cuál será el pago en este caso según la recta de regresión? Justifica si es razonable realizar dicha predicción.
- (2 puntos) Representa un diagrama de caja para los datos de la variable *Salarios* y determina si existen datos atípicos. Justifica la respuesta.

Problema 2 (10 puntos) Los autobuses urbanos de cierta ciudad tardan en realizar sus trayectos un tiempo X que puede modelizarse según una distribución uniforme en el intervalo (30, 40).

- (1 punto) Dibuja la función de densidad de la variable aleatoria X . Especifica en los ejes de coordenadas tanto los valores de X con densidad positiva, como los valores que toma la función de densidad.
- (2 puntos) ¿Cuál es la probabilidad de que un autobús tarde entre 30 y 37.5 minutos en realizar su trayecto?
- (3 puntos) Se seleccionan aleatoriamente 100 autobuses urbanos de esta ciudad y se cuenta cuántos de ellos tardarán entre 30 y 37.5 minutos en realizar sus trayectos. Especifica el nombre de la distribución de esta nueva variable aleatoria, que llamaremos Y , e indica los valores de sus parámetros. Calcula la esperanza y desviación típica de Y .
- (4 puntos) ¿Cuál es la probabilidad (aproximada) de que haya menos de 64 autobuses que tarden entre 30 y 37.5 minutos en realizar sus trayectos?

Problema 3 (10 puntos) Según un estudio sobre los hábitos de consumo de las familias de la Comunidad de Madrid, el gasto mensual en productos de marcas blancas de las familias puede modelizarse según una variable aleatoria X con ley Normal de desviación típica 50 euros. Los datos siguientes corresponden a una muestra aleatoria simple de tamaño $n = 10$ familias:

<i>Familia</i>	1	2	3	4	5	6	7	8	9	10
<i>Gasto</i>	325	193	203	305	251	240	233	238	354	311

- (2.5 puntos) Define qué es un estimador insesgado (o centrado) y propón uno para el gasto medio mensual en productos de marcas blancas de las familias de la Comunidad de Madrid, razonando tu respuesta. ¿Qué estimación del gasto medio mensual se obtiene para el estimador propuesto?
- (2.5 puntos) Compara el estimador propuesto en el apartado a) con el siguiente:

$$T = \frac{X_1 + X_n}{2},$$

donde X_1, X_2, \dots, X_n es una muestra aleatoria simple de X . Justifica con alguno de los criterios que conoces con cuál de ellos te quedarías. ¿Qué estimación se obtiene para el estimador T ?

- (2.5 puntos) Calcula un intervalo de confianza al 95% para el gasto medio mensual en productos de marcas blancas de las familias de la Comunidad de Madrid, especificando las hipótesis necesarias para calcularlo.
- (2.5 puntos) Si queremos trabajar con niveles de confianza del 90%, ¿obtendríamos un intervalo más o menos amplio que el calculado anteriormente? Razona tu respuesta sin calcular el nuevo intervalo de confianza.