

Lección 4. Ejercicios complementarios.

Ejercicio 1 (junio 06). La información relativa al mes de enero sobre los ingresos (X) y los gastos (Y), expresados en miles de euros, de un conjunto de pequeñas empresas, se ha utilizado para estimar la recta de regresión que, para dicho conjunto de empresas, explica los gastos del mes en función de los ingresos. Si la recta estimada es:

$$y^* = 0,15 + 1,2x$$

y se sabe, además, que:

$$\bar{x} = 7,5; \quad s_x^2 = 9,2 \quad y \quad s_y^2 = 14,2,$$

obtenga:

- Una estimación del gasto en el mes de enero de una empresa cuyos ingresos mensuales fueron de 8.000 euros.
- ¿Qué fiabilidad le merece dicha predicción? Justifique su respuesta.
- ¿Cuál es el gasto medio del mes por empresa?
- ¿Qué porcentaje de las variaciones en los gastos mensuales de esas empresas no vienen explicadas por las variaciones en los ingresos mensuales de las mismas a través del modelo lineal?

Ejercicio 2 (Diciembre 06). Utilizando los datos mensuales correspondientes al año 2005, se desea estimar un modelo lineal que explique el Grado de Ocupación hotelera en Andalucía (en %) en función del número de pernoctaciones (en millones). La información resumida sobre las variables X = "Número de pernoctaciones, en millones" e Y = "Grado de ocupación hotelera, en porcentaje" se muestra a continuación:

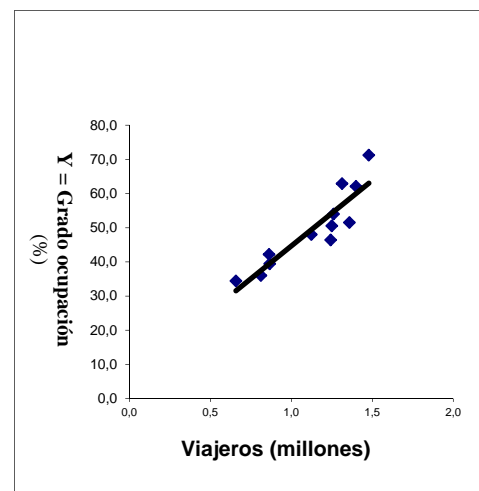
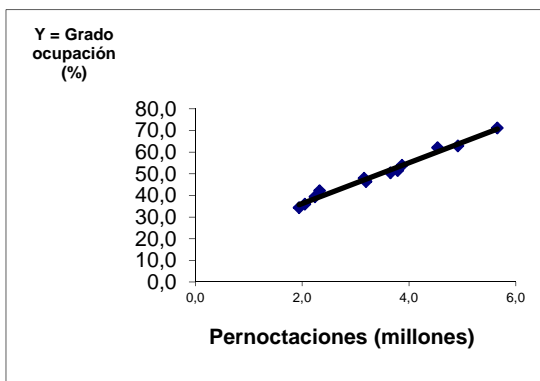
$$\sum Y_i = 598,6; \quad \sum X_i = 41,3; \quad \sum Y_i X_i = 2209,5; \quad \sum Y_i^2 = 31270,3; \quad \sum X_i^2 = 158,1,$$

Nota: Los datos originales proceden de la Encuesta de Ocupación en alojamientos hoteleros (EOH). INE.

Se pide:

- Estime la recta de regresión que explica el Grado de Ocupación Hotelera en función de las pernoctaciones.
- Calcule el Coeficiente de Determinación (R^2) del ajuste anterior e interprete el resultado.
- Sabiendo que el número de pernoctaciones en el mes de enero de 2006 fue 2.200.000, utilice el modelo anterior para estimar el Grado de Ocupación Hotelera en Andalucía en dicho mes y comente la fiabilidad de dicha predicción.

Ejercicio 3 (Febrero 06).- Una importante empresa multinacional del sector turístico desea explicar, mediante un modelo lineal, el comportamiento del Grado de Ocupación por plazas en Andalucía (%) en los doce meses del año 2005. Con este fin, se plantea utilizar las observaciones mensuales correspondientes a dos posibles variables explicativas: número mensual de pernoctaciones (en millones) o el número mensual de viajeros (en millones de personas). Los gráficos correspondientes a las nubes de puntos y a las dos posibles líneas de regresión figuran a continuación:



Se sabe que:

$$\sum Y_i = 598,6; \quad \sum X_i = 41,3; \quad \sum Y_i X_i = 2209,5; \quad \sum Y_i^2 = 31270,3; \quad \sum X_i^2 = 158,1,$$

donde X son las pernoctaciones en millones. Por otra parte, en cuanto a la recta de regresión que explica el Grado de ocupación hotelera (Y) en función

del número de viajeros, se han hecho los cálculos correspondientes y se sabe que:

- i) El Grado de Ocupación Hotelera asciende a 6,3% cuando el número de viajeros es cero.
- ii) Cuando el número de viajeros aumenta en un millón de personas el Grado de Ocupación Hotelera aumenta en 38,4 puntos porcentuales.
- iii) El 17,8% de las variaciones en el Grado de Ocupación Hotelera no vienen explicadas por el número de viajeros a través del modelo lineal.

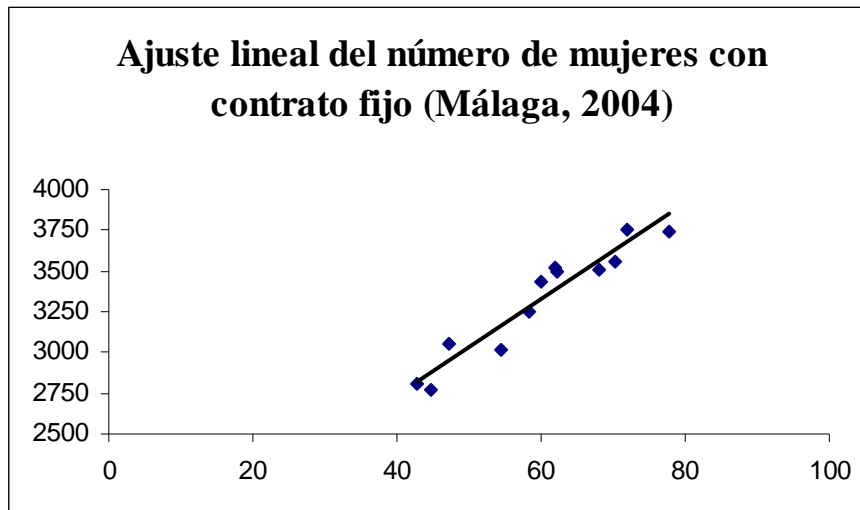
Nota: Los datos originales proceden de la Encuesta de Ocupación en alojamientos hoteleros (EOH). INE.

Se pide:

- a) Usando sólo la información de los gráficos, ¿qué variable cree usted que podría explicar mejor el Grado de Ocupación Hotelera? Justifique su respuesta.
- b) Estime la recta de regresión que explica el Grado de Ocupación Hotelera en función de las pernoctaciones.
- c) Utilice la información de los puntos i) al iii), para determinar los parámetros de la recta que explica el Grado de Ocupación Hotelera en función del número de viajeros.
- d) Se sabe que el Coeficiente de Determinación (R^2) de la recta de regresión que explica el Grado de Ocupación Hotelera en función de las pernoctaciones asciende a 0,994. ¿Cuál es el valor del Coeficiente de Determinación (R^2) de la otra recta de regresión?
- e) Teniendo en cuenta el apartado anterior y sabiendo que:
 - El número de pernoctaciones para el mes de enero de 2006 asciende a 2.200.000 pernoctaciones
 - El número de viajeros para el mes de enero de 2006 asciende a 800.000 personas,

obtenga una predicción del Grado de Ocupación Hotelera en el mes de enero de 2006 utilizando el modelo más adecuado.

Ejercicio 4 (Septiembre 06). Con los datos mensuales del año 2004 se ha estimado un modelo de regresión lineal que explica el número de mujeres con contrato fijo en el sector de la Hostelería en la provincia de Málaga, Y_i , en función de la tasa de ocupación hotelera de dicha provincia en porcentaje, X_i . En el siguiente gráfico se muestra la nube de puntos junto a la recta ajustada.



Se sabe que:

$$\sum Y_i = 39.923 ; \quad \sum X_i = 719,6 ; \quad \sum Y_i X_i = 2.434.211,9 ; \quad \sum Y_i^2 = 134.107.283 ; \quad \sum X_i^2 = 44.523.$$

- Estime la recta de regresión que explica el número de mujeres con contrato fijo en función del grado de ocupación.
- Calcule la bondad del ajuste e interprete el resultado.
- Obtenga una predicción del número de mujeres con contrato fijo para el mes de enero de 2005 suponiendo que el grado de ocupación hotelera ha sido del 52%.

Ejercicio 5 (Febrero 07). Haciendo uso de la Encuesta de Ocupación Hotelera (EOH) con datos anuales desde 1999 hasta 2006, se desea estudiar la relación existente entre el número de viajeros en Andalucía (millones de personas) que figura la tabla adjunta con el nombre de V_AND y el número de Viajeros en España (millones de personas) que figura con el nombre de V_ESP. En la siguiente tabla aparecen sólo los cuadrados de ambas variables:

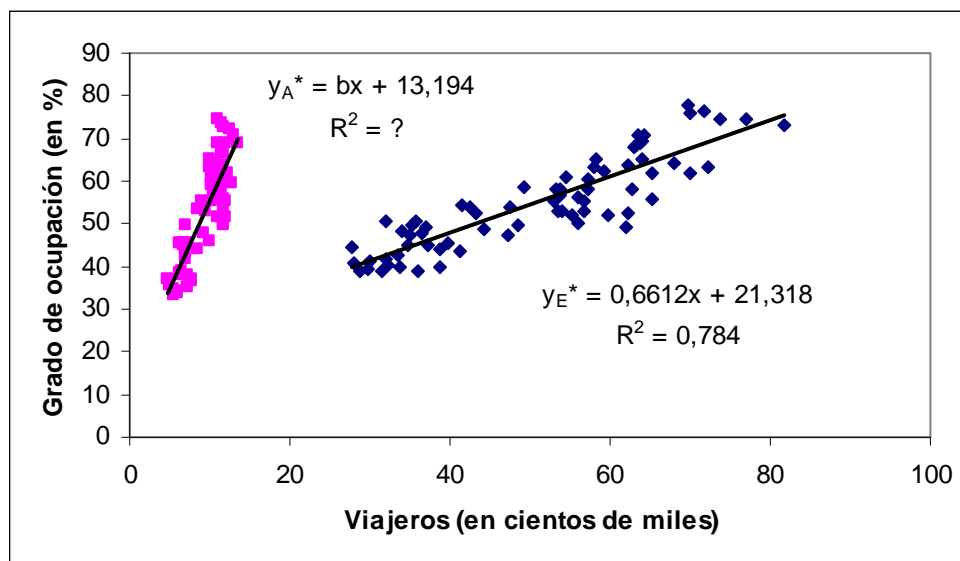
	V_AND ²	V_ESP ²
1999	113,9	3.432,7
2000	124,7	3.514,4
2001	127,9	3.588,5
2002	130,1	3.584,3
2003	140,5	3.910,2
2004	157,1	4.466,4
2005	185,8	4.988,5
2006	241,1	6.729,2
Total	1.221,07	34.214,12

Se sabe además que, para el periodo considerado:

- El número medio de viajeros al año en Andalucía fue de 12,26 millones de personas.
 - El número medio de viajeros al año en España fue de 64,96 millones de personas.
 - El coeficiente de correlación lineal simple entre las dos variables ascendió a 0,9937.
- a) Estime una recta de regresión que permita explicar el comportamiento del número anual de viajeros en Andalucía a partir del correspondiente al número anual de viajeros en España.
 - b) Explique el significado, matemático y económico, de los parámetros estimados en el apartado anterior.
 - c) ¿Qué porcentaje de las variaciones en el número anual de viajeros en Andalucía no vienen explicadas por las variaciones en el número anual de viajeros en España a través del modelo lineal estimado?
 - d) ¿A cuánto asciende el error cometido en la estimación, a partir de la recta de regresión, del número de viajeros en Andalucía en el año 2006?
 - e) Obtenga una predicción del número de viajeros en Andalucía para el año 2010, suponiendo que el número de viajeros en España para ese año fuese de 62.530.000 personas. Comente la fiabilidad de la predicción.

Ejercicio 6 (Septiembre 07). Con los 72 datos mensuales anteriores a Enero de 2005 que, sobre viajeros y grado de ocupación por plazas, para Andalucía y España, ofrece la Encuesta de Ocupación Hotelera (EOH) que elabora el INE, se han estimado las rectas de regresión mínimo-cuadráticas para explicar, mediante un modelo lineal, las variaciones en el grado de ocupación a partir de las variaciones en el número de viajeros. En el Gráfico 1 se presentan las nubes de puntos y las rectas ajustadas (y_A^* e y_E^*) para Andalucía y España, respectivamente.

GRÁFICO 1



Se pide:

- Indique el significado, tanto matemático como en el contexto del enunciado, de los parámetros de la recta estimada para España, $y_E^* = 0,6612x + 21,318$.
- Indique el porcentaje de las variaciones en el grado de ocupación hotelera en España que no vienen explicadas por las variaciones en el número de viajeros a través del modelo lineal estimado.
- Si en el periodo utilizado para la estimación, por término medio, mensualmente visitaron España 5.097.323 viajeros, ¿cuál fue, por término medio, el grado de ocupación mensual en España durante ese periodo?
- Determine, haciendo uso de los datos auxiliares que se le facilitan a continuación, la pendiente de la recta estimada para Andalucía y el coeficiente de determinación de dicho ajuste sabiendo que el grado de ocupación hotelera viene expresado en tantos por ciento y el número de viajeros en cientos de miles.

$$\sum_i y_i = 3855,85; \sum_i x_i = 689,41626; \sum_i y_i^2 = 215.900,559; \sum_i x_i^2 = 6.992,45437; \sum_i x_i y_i = 38.569,2725$$

- e) Estime el grado de ocupación hotelera en Enero de 2005, tanto para España como para Andalucía, teniendo en cuenta que el número de viajeros en Enero de 2005 en España y Andalucía ascendió a 3.356.718 y 656.711 viajeros, respectivamente.

Ejercicio 7 (Febrero 08). Un analista se plantea utilizar la técnica de regresión lineal para predecir el Índice de Precios en Hoteles, cafés y restaurantes en Andalucía, Y, en función del Índice de Precios en Hoteles, cafés y restaurantes en España (X). Si a partir de los datos trimestrales desde 2002 hasta 2007 ha obtenido la siguiente información:

$$\sum_{i=1}^{24} Y_i = 2.246,05; \sum_{i=1}^{24} X_i = 2.261,563; \sum_{i=1}^{24} X_i Y_i = 212.891,292; \sum_{i=1}^{24} X_i^2 = 214.269,612;$$
$$\sum_{i=1}^{24} Y_i^2 = 211.530,04,$$

determine:

- La recta de regresión indicada, interpretando el significado de los parámetros estimados.
- El porcentaje de las variaciones en el Índice de Precios en Andalucía no vienen explicadas por las variaciones en el Índice de Precios en España a través de la recta de regresión.
- Una predicción del Índice de Precios en Andalucía, suponiendo que el Índice de Precios en España asciende a 106,8.