Ejercicios de Autocomprobación

Tema 6

1. En una investigación sobre la relación entre el tráfico X (en miles de automóviles cada 24 horas) y el contenido en plomo Y de la corteza de los árboles cerca de la autopista (en $\mu g/g$ de peso seco), se obtuvieron los datos de la tabla siguiente:

x	8,3	8,3	12,1	12,1	17	17	17	24,3	24,3	24,3	33,6
Y	227	312	362	521	640	539	728	945	738	759	1263

   a) Calcular la relación inicial que permita predecir, a partir del tráfico, el contenido en plomo en la corteza de un árbol. ¿Cuál es el coeficiente de correlación lineal?
   b) Contrastar al 90% la bondad del modelo.
   c) Contrastar la hipótesis de que un aumento en un punto en el tráfico provoca un incremento de 40 unidades en el contenido en plomo en la corteza de los árboles.
   d) Con un 95% de confianza, ¿entre qué valores oscilará el contenido en plomo de la corteza de los árboles cerca de una autopista con un tráfico de 20000 automóviles cada 24 horas?

2. Un gran almacén recopiló información sobre 10 campañas de ventas, de los gastos en publicidad (X) y los beneficios netos (Y). Los resultados obtenidos son los siguientes (ambas variables medidas en miles de euros):

$$\sum_{i=1}^{10} x_i = 693, \quad \sum_{i=1}^{10} x_i^2 = 48141, \quad \sum_{i=1}^{10} x_i y_i = 10948,$$

$$\sum_{i=1}^{10} y_i = 1575, \quad \sum_{i=1}^{10} y_i^2 = 254245$$

   a) Calcular el coeficiente de correlación y la recta de regresión que permita predecir los beneficios netos en función de los gastos de publicidad.
   b) Obtener un intervalo de confianza al 90% para la pendiente del modelo. Utilizando el intervalo obtenido, contrastar la afirmación de la agencia de publicidad de que mil euros gastados en publicidad aporta más de cinco mil euros de beneficios.

3. Una empresa inmobiliaria estudia la relación entre el tiempo en semanas que se tarda en vender los pisos (Y) y el precio, en miles de euros, que se pide por ellos (X). La siguiente tabla contiene los datos obtenidos en una muestra de diez pisos.

x	220	275	300	242	256	225	155	175	305	195
Y	5,5	12	11,5	11	8,6	8,5	5,2	4,8	14,2	7

a) Obtener la recta de regresión que permita predecir el tiempo que se tarda en vender un piso en función de su precio.
b) Contrastar que la pendiente de la recta es distinta de cero.
c) Predecir el tiempo que se tarda en vender una casa de 200.000 euros, dando un intervalo de confianza del 90%.

4. Una compañía quiere estudiar la relación entre el número de máquinas que esperan revisión en un momento determinado y el tiempo promedio que necesitan los operadores para servir las máquinas. La compañía elige al azar ocho registros que muestran el número de máquinas en espera al comienzo de un periodo dado (X) y el número de servicios empleados por el operario durante el periodo (Y). Los resultados obtenidos son los siguientes:

$$\sum_{i=1}^{8} x_i = 43, \quad \sum_{i=1}^{8} x_i^2 = 251, \quad \sum_{i=1}^{8} x_i y_i = 180,$$
$$\sum_{i=1}^{8} y_i = 32, \quad \sum_{i=1}^{8} y_i^2 = 144$$

a) Calcular la recta de regresión que permita predecir el número de servicios que cumplirá un operario a partir del número de máquinas en espera al principio del periodo.
b) Obtener un intervalo de confianza al 90% para la pendiente del modelo.

5. En un estudio, hecho por el departamento de transporte de Sin-City, sobre el efecto de los precios del billete de autobús (en céntimo de euro) sobre el número de pasajeros (en miles), se obtuvieron los resultados de la tabla:

Precio del billete	25	30	35	40	45	55	60
Miles de viajeros	800	780	780	660	600	620	620

a) Obtener la recta de regresión que permita predecir el número de pasajeros (Y) a partir del precio del billete (X) y calcular el coeficiente de correlación.
b) Contrastar la afirmación de que el número de miles de pasajeros que se pierden al aumentar en un céntimo el precio del billete es menor o igual a 5. Utilizar $\alpha = 0{,}05$. Calcular además el nivel crítico o p-valor aproximado.

Resultado numérico de los ejercicios de autocomprobación

1. a) La recta de regresión es $\hat{Y}_i = 36{,}18x_i - 12{,}7668$. El coeficiente de correlación lineal es $r = 0{,}956$.

   b) Se rechaza la hipótesis nula y por tanto el coeficiente β_1 es significativamente distinto de cero y el modelo es lineal.

   c) Se obtiene un nivel crítico o p-valor $p = 0{,}3396$ y no se rechaza la hipótesis nula.

   d) Como $x_h = 20$, el intervalo al 95% de confianza para la predicción es $(492{,}3012; 929{,}3652)$.

2. a) El coeficiente de correlación lineal es $r = 0{,}9448$. La recta de regresión es $\hat{Y}_i = 6{,}8949x_i - 320{,}31$.

   b) El intervalo de confianza al 90% para β_1 es $(5{,}3233; 8{,}4665)$. Como el intervalo obtenido está a la derecha del valor 5, no se puede rechazar la afirmación de la agencia de publicidad (en base a la muestra observada).

3. a) La recta de regresión es $\hat{Y}_i = 0{,}058x_i + 1{,}038576$.

   b) Se obtiene un nivel crítico o p-valor inferior a $0{,}01$ y se rechaza la hipótesis nula de que la pendiente es cero.

   c) Como $x_h = 200$, el intervalo al 90% de confianza para la predicción es $(3{,}95; 9{,}67)$.

4. a) La recta de regresión es $\hat{Y}_i = 0{,}4025x_i + 1{,}8365$.

   b) El intervalo de confianza al 90% para β_1 es $(-0{,}233576; 1{,}038576)$.

5. a) La recta de regresión es $\hat{Y}_i = -6{,}238x_i + 952{,}62$. El coeficiente de correlación lineal es $r = -0{,}908$.

   b) La hipótesis nula es $\beta_1 = -5$ y la hipótesis alternativa es $\beta_1 < -5$. Al 95% de confianza no se rechaza la hipótesis nula. El p-valor es $p = 0{,}173$.