$\frac{\quad}{4} + \frac{\quad}{6} = \frac{\quad}{10}$ CALIFICACIÓN:

El gerente de una empresa de artículos de deportes con tiendas en Francia, Italia y Reino Unido, quiere analizar los factores que influyen en la determinación de los ingresos de sus tiendas. Para ello dispone de una muestra correspondiente a 265 tiendas sobre las siguientes variables

I_i: Ingresos anuales de la tienda i-ésima (miles de euros).

GP_i: Gastos en publicidad de la tienda i-ésima (miles de euros).

R_i: Renta promedio del área donde se encuentra la tienda i-ésima (miles de euros).

PARTE 1 (4 puntos)

Se especifica el siguiente modelo de regresión lineal general para determinar los ingresos:

$$(1) \qquad I_i = \beta_0 + \beta_1 GP_i + \beta_2 R_i + u_i \qquad i = 1, 2, \ldots, 265$$

Los resultados de su estimación por MCO son:

$$\hat{I}_i = -7,881 + \underset{(3,47)}{17,491} GP_i + \underset{(0,057)}{0,306} \; R_i \qquad R^2 = 0,4793 \quad SCR = 234562$$

1. Manteniéndose el resto de los factores constantes, ¿en cuántos euros estimas que aumentan los ingresos de una tienda si los gastos en publicidad aumentan en cinco mil euros?

   $\hat{\beta}_1 \times 5 = 17,91 \times 5 = 89,55$ miles de euros, es decir, 89550 euros.

   Esto se debe a que el coeficiente β_1 mide la variación esperada en los ingresos ante un aumento de mil euros en los gastos de publicidad manteniendo constante la renta.

2. Interpreta el valor del coeficiente de determinación.

   El 47,9 % de la variabilidad de los ingresos de las tiendas en la muestra viene explicada por la variabilidad de las variables explicativas gastos en publicidad y renta en términos lineales.

3. Contrasta la significación conjunta de las variables explicativas.

   $H_o : \beta_1 = \beta_2 = 0$
   $H_a : \beta_1 \neq 0$ y/o $\beta_2 \neq 0$

   $$F = \frac{R^2/q}{(1-R^2)/(T-(k+1))} \overset{H_o}{\approx} \mathcal{F}(q, T-(k+1))$$

   $$F = \frac{0,4799/2}{(1-0,5299)/(265-3)} = 120,44 > 3,00 \approx \mathcal{F}_{0,05}(2, 262)$$

   Se rechaza la hipótesis nula con un nivel de significación del 5 % por lo que se puede concluir que las variables explicativas gastos en publicidad y renta son conjuntamente significativas.

4. En base a la recta de regresión muestral, estima cuánto debería gastarse en publicidad una tienda situada en un barrio con una renta media de 25000 euros si quiere conseguir unos ingresos de 1 millón de euros.

$$\hat{I}_i = -7,881 + 17,491 GP_i + 0,306 R_i$$

$$1000 = -7,881 + 17,491 GP_i + 0,306 \times 25 \quad \Rightarrow \quad \widehat{GP}_i = 55,821 \text{ miles de euros (55821 euros)}$$

PARTE 2 (6 puntos)

La muestra disponible cuenta con ingresos de tiendas correspondientes a los tres países citados: Francia, Italia y Reino Unido. Las primeras 90 observaciones corresponden a tiendas ubicadas en Francia, las siguientes 87 observaciones a tiendas situadas en Italia y las restantes 88 se encuentran en el Reino Unido. Debido a esto, el gerente mantiene la teoría de que existen diferencias en los ingresos de las tiendas según sea su localización.

1. Escribe un MRLG que incluya la nueva variable explicativa localización en este modelo, explicando detalladamente su especificación.

   La Localización es una variable explicativa cualitativa con tres categorías: la tienda está en Francia o está en Italia o está en Reino Unido. Por lo tanto, para introducir este nuevo regresor en el modelo se han de definir tres variables ficticias, una por cada categoría:

   $$\text{Francia}_i = \begin{cases} 1 & i \in \text{ tienda en Francia} \\ 0 & \text{en otro caso} \end{cases} \qquad \text{Italia}_i = \begin{cases} 1 & i \in \text{ tiendas en Italia} \\ 0 & \text{en otro caso} \end{cases}$$

   $$\text{RUnido}_i = \begin{cases} 1 & i \in \text{ tiendas en Reino Unido} \\ 0 & \text{en otro caso} \end{cases}$$

   Al especificar el modelo hay que tener cuidado para no tener problemas de colinealidad perfecta. Se especifica el modelo incluyendo tantas variables ficticias como categorías menos 1 y con término constante. Dejando fuera la variable ficticia $Italia_i$ el modelo es:

   $$(2) \qquad I_i = \beta_0 + \beta_1 \, GP_i + \beta_2 \, R_t + \beta_3 \, \text{Francia}_i + \beta_4 \, \text{RUnido}_i + u_i$$

2. El resultado de estimar el modelo propuesto en el apartado anterior es:

   $$\hat{I}_i = -13,71 + \underset{(3,38)}{15,62} \, GP_i + \underset{(0,055)}{0,286} \, R_i + \underset{(4,66)}{19,751} \, \text{Francia}_i + \underset{(4,72)}{13,50} \, \text{RUnido}_i \qquad SCR = 223746$$

   Interpreta los coeficientes estimados que acompañan a las variables Francia y RUnido.

   $\hat{\beta}_3 = 19,751$: diferencia estimada en los ingresos de las tiendas de Francia con respecto a las de Italia, manteniendo el resto de las variables explicativas (gastos en publicidad y renta) constantes.

   $\hat{\beta}_4 = 13,495$: diferencia estimada en los ingresos de las tiendas de Reino Unido con respecto a las de Italia, manteniendo el resto de las variables explicativas (gastos en publicidad y renta) constantes.

3. Estima los ingresos de las siguientes tiendas:

   - Observación 1 Gastos en Publicidad 10000 euros Renta 20000 euros
   - Observación 100 Gastos en Publicidad 20000 euros Renta 35600 euros

$$i = 1 \qquad \hat{I}_i = -13,71 + 15,62 \times 10 + 0,286 \times 20 + 19,751 \times 1 + 13,50 \times 0 = 167,961$$

$$i = 100 \qquad \hat{I}_i = -13,71 + 15,62 \times 20 + 0,286 \times 35,6 + 19,751 \times 0 + 13,50 \times 0 = 308,87$$

4. Contrasta la hipótesis del gerente de que la variable explicativa localización influye en los ingresos de las tiendas.

   $H_o : \beta_3 = \beta_4 = 0$
   $H_a : \beta_3 \neq 0$ y/o $\beta_4 \neq 0$

   Estadístico de contraste:

$$F = \frac{(SCR_R - SCR_{NR})/q}{SCR_{NR}/(T-(k+1))} \overset{H_o}{\approx} \mathcal{F}(q, T-(k+1))$$

   donde SCR_{NR} es la suma de cuadrados de residuos del modelo no restringido (Modelo 2) y SCR_R es la suma de cuadrados de residuos del modelo restringido (Modelo 1).

$$F = \frac{(234562,00 - 223746,00)/2}{223746,00/(265-5)} = 6,284 > \mathcal{F}(2,260)_{0,05} \approx 3$$

   Se rechaza la hipótesis nula con un nivel de significación del 5 %, es decir, concluimos que la variable localización es significativa.

5. De acuerdo a los resultados previos obtenidos, ¿qué modelo elegirías para determinar los ingresos?, ¿por qué? Razona tu respuesta y explica cuáles son las propiedades de los estimadores en el modelo que NO has elegido.

   En el modelo (2) todas las variables son significativas individualmente: todas las variables que ya estaban incluidas en el Modelo 2 tienen estadísticos-t superiores al valor crítico. Además, en el apartado anterior se acaba de comprobar que la variable localización también es individualmente significativa. Por lo tanto, elegiremos el modelo (2) porque el modelo (1) está dejando fuera una variable explicativa relevante, la localización.

   Como en el modelo (1) se omiten variables relevantes, los estimadores MCO están sesgados. Además, también estará sesgada la estimación de la varianza de las perturbaciones, lo que llevará a calcular estadísticos t y F que no siguen la distribución propuesta y por lo tanto invalidará los contrastes. Todo ello lleva a elegir el modelo (2) que, a priori, no muestra indicios de mala especificación.

6. A pesar de todo, el gerente se sigue cuestionando si el modelo recoge adecuadamente la influencia de la localización sobre los ingresos. Piensa que el efecto de los gastos en publicidad sobre los ingresos puede ser diferente por países. Especifica un modelo que recoja este efecto, explicándolo detalladamente.

   En los modelos (1) y (2) la variación esperada en el ingreso ante cambios en los gastos en publicidad es constante independientemente del país. Para permitir que el efecto marginal de los gastos en publicidad sobre los ingresos sea diferente para cada país hay que introducir un efecto interacción entre las dos variables, gastos en publicidad y localización ($GP \times$ localización).

   Dado que la variable localización es cualitativa hay que multiplicar la variable gastos en publicidad por cada una de las variables ficticias que la representa:

$$(GP \times \text{Francia})_i = \begin{cases} GP_i & i \in \text{ en Francia} \\ 0 & \text{en otro caso} \end{cases} \qquad (GP \times \text{Italia})_i = \begin{cases} GP_i & i \in \text{ en Italia} \\ 0 & \text{en otro caso} \end{cases}$$

$$(GP \times \text{RUnido})_i = \begin{cases} GP_i & i \in \text{ en Reino Unido} \\ 0 & \text{en otro caso} \end{cases}$$

   Al especificar el modelo hay que tener cuidado para no tener problemas de colinealidad perfecta. Se especifica el modelo incluyendo tantas variables como interacciones menos 1. Dejando fuera la interacción $(GP \times \text{Italia})_i$:

$$I_i = \beta_0 + \beta_1 \, GP_i + \beta_2 \, R_t + \beta_3 \, \text{Francia}_i + \beta_4 \, \text{RUnido}_i + \beta_5 \, (GP \times \text{Francia})_i + \beta_6 \, (GP \times \text{RUnido})_i + u_i$$

   En este modelo el efecto marginal de los gastos en publicidad sobre los ingresos, dada la renta es:

$$\frac{\partial E(I_i)}{\partial GP_i} = \beta_1 + \beta_5 \, \text{Francia}_i + \beta_6 \, \text{RUnido}_i \qquad \begin{cases} \text{Francia} & \frac{\partial E(I_i)}{\partial GP_i} = \beta_1 + \beta_5 \, \text{Francia}_i \\ \text{Reino Unido} & \frac{\partial E(I_i)}{\partial GP_i} = \beta_1 + \beta_6 \, \text{RUnido}_i \\ \text{Italia} & \frac{\partial E(I_i)}{\partial GP_i} = \beta_1 \end{cases}$$

β_1: variación esperada en los ingresos si los gastos en publicidad aumentan en mil euros y la tienda está en Italia, manteniendo la renta constante.

β_5: diferencia en la variación esperada en los ingresos al aumentar los gastos en publicidad en mil euros entre una tienda que está en Francia con respecto a una que esté en Italia, manteniendo la renta constante.

β_6: diferencia en la variación esperada en los ingresos al aumentar los gastos en publicidad en mil euros entre una tienda que está en Reino Unido con respecto a una que esté en Italia, manteniendo la renta constante.