$\overline{11}$	+	$\overline{4}$	+	$\overline{8}$	=	$\overline{23}$	CALIFICACION:	

Uno de los estudios econométricos más conocidos y citados en la literatura sobre el impacto a largo plazo de la publicidad es el realizado en 1963 por K. Palda para su tesis doctoral[1].

Este investigador analizó la Lydia E. Pinkham Medicine Company, que ha manufacturado desde 1873 una medicina patentada que se suponía que no sólo aliviaba el dolor sino que también era capaz de curar una amplia gama de enfermedades. Se considera que los datos de esta empresa forman un conjunto ideal para analizar los efectos de la publicidad en las ventas por varias razones. En primer lugar, su producto no tenía que enfrentarse a publicidad de productos rivales y, además, la publicidad era el único elemento de marketing que utilizaba la empresa. En particular, el producto de L.E.P no tenía un rival directo, no utilizaba vendedores o agentes de ventas y no tenía un sistema de distribución digno de mención salvo por la publicidad. En segundo lugar, L.E.P. era uno de los mayores anunciantes a nivel de Estados Unidos y bastante controvertido. La ratio Publicidad/Ventas era excepcionalmente alta, llegando al 85 % en 1934 y bajando del 40 % en muy raras ocasiones. La controversia surge del hecho de que este producto consistía en un extracto de hierbas en una solución alcohólica. En 1914, el hecho de que este producto "analgésico" contenía un 18 % de alcohol, indujo a la Hacienda estadounidense (Internal Revenue Service) a amenazar con hacerle pagar impuestos como si de una bebida alcohólica se tratara y al Ministerio de Consumo (Food and Drug Administration) a acusar a la compañía de publicidad falsa y engañosa. En respuesta, la compañía redujo severamente en su publicidad las pretensiones curativas de su producto y rebajó el contenido de alcohol a un 15 %. Pero, las medidas tomadas por el Ministerio de Consumo en 1925, obligaron a L.E.P. a anunciar el producto como un tónico de hierbas, con lo que las ventas cayeron de forma inmediata. En 1926, Lydia Gove, la nieta de la fundadora fue nombrada gerente de la compañía y se empeñó en no cortar los gastos en publicidad aunque las ventas bajaban enfrentandose al resto del consejo de dirección. Trás un duro pleito, la sra. Gove fue retirada del cargo en 1937, y los gastos en publicidad de la compañía no sólo se redujeron drásticamente sino que además cambiaron de medio, es decir, de revistas y periódicos se pasó mayoritariamente a la radio.

Por estas dos razones se considera que los datos de esta compañía son relativamente "limpios". El problema que presentan estos datos, sin embargo, es que el precio no se puede utilizar ni como variable explicativa ni como deflactor. Por un lado, no hay datos sobre los precios al por menor y los precios al por mayor cambian poco y además sus cambios a veces no están relacionados con los cambios de precios al por menor. Por otro lado, hay datos sobre el valor de las ventas y sobre su volumen (número de unidades físicas), pero como el producto cambia a lo largo del tiempo de ser un tónico líquido a tabletas, tampoco se puede utilizar esta información para obtener el precio unitario. Por lo tanto, habrá que ignorar el efecto de los precios sobre la demanda. Además, como tampoco se puede encontrar un deflactor para la publicidad, se trabaja con ventas y gastos en publicidad en dólares corrientes.

Utilizaremos estos datos (fichero `lidia.gdt`) para analizar el efecto a largo plazo de la publicidad sobre las ventas de un producto.

[1]Publicado en Palda, K. (1964). *The Measurement of Cumulative Advertising Effects*, Englewoods Cliffs, N.J. Prentice Hall.

PARTE 1 (11 puntos)

Comenzaremos por plantear un modelo de regresión lineal simple que relacione las ventas del producto con los gastos en publicidad (ambas medidas en miles de dólares):

$$V_t = \beta_0 + \beta_1 \, GP_t + u_t \qquad t = 1, 2, \ldots, 54$$

El fichero `lidia.gdt` contiene más variables pero, por ahora, consideraremos sólo las ventas y los gastos en publicidad.

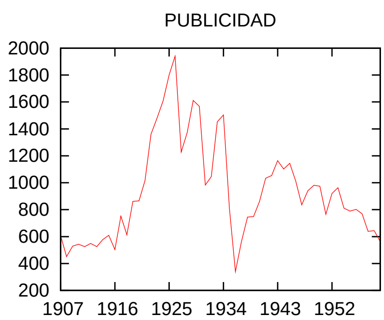
1. Describe los datos para las dos variables de interés: frecuencia, duración, estadísticos descriptivos.

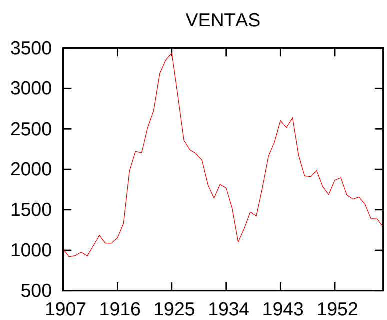
   Los datos son anuales de 1907 a 1960 y los principales estadísticos descriptivos se encuentran en la siguiente tabla:

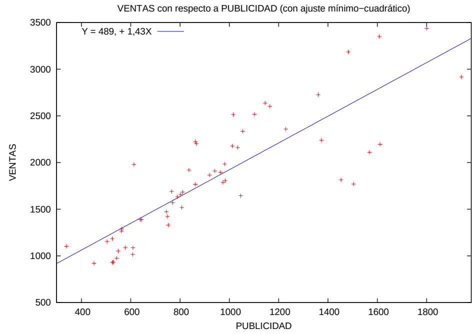
Variable	Media	Mediana	Mínimo	Máximo
PUBLICIDAD	934,519	862,000	339,000	1941,00
VENTAS	1829,48	1778,50	921,000	3438,00

Variable	Desv. Típ.	C.V.	Asimetría	Exc. de curtosis
PUBLICIDAD	371,879	0,397936	0,806414	−0,0181078
VENTAS	632,738	0,345857	0,602604	−0,129694

2. Representa gráficamente cada variable por separado, así como su scatterplot. ¿Qué te sugiere este último en términos de la relación existente entre las dos variables?







   En el gráfico Ventas-Gastos de publicidad se observa que las ventas creen con la publicidad, por lo tanto, la relación entre ambas es positiva. Además, el crecimiento de las ventas es más fuerte para niveles de publicidad más bajos mientras que para los niveles más altos de publicidad, los resultados en las ventas son más volátiles. La relación lineal entre ambas variables parece razonable pero también podría plantearse la conveniencia de una relación no lineal tipo logarítmica. Esta forma funcional implica que las ventas crecen al aumentar la publicidad pero la tasa de crecimiento decrece a mayor nivel de gastos en publicidad.

3. Estima el modelo por Mínimos Cuadrados Ordinarios e interpreta los resultados:

   - Escribe la recta de regresión muestral.

$$\hat{V}_t = \underset{(127,44)}{488,83} + \underset{(0,13)}{1,43}\; GP_t \qquad R^2 = 0,71 \quad SCR = 6134381 \qquad t = 1907, \ldots, 1960$$

   - Interpreta los parámetros estimados.
     - $\hat{\beta}_0$: ventas estimadas en miles de dólares cuando los gastos en publicidad son cero.
     - $\hat{\beta}_1$: variación estimada en el volumen de ventas (en miles de dólares) cuando los gastos en publicidad aumentan en mil dólares.

   - ¿Cuál es la bondad de ajuste del modelo? Interpreta el estadístico correspondiente. El coeficiente de determinación $R^2 = 0,71$ mide la bondad de ajuste del modelo: el 71 % de la variabilidad de las ventas en la muestra queda explicada por la variabilidad en los gastos en publicidad en términos lineales.

4. ¿Qué consecuencias tendría sobre la estimación del modelo medir las variables en millones de dólares? Demuestra tus conclusiones empíricamente.

   Para medir las variables en millones de dólares hay que dividirlas por mil:

$$V_t^* = V_t/1000 \qquad\qquad GP_t^* = GP_t/1000$$

   Por lo tanto:

$$\bar{V}_t^* = \bar{V}/1000 \qquad\qquad \bar{GP}^* = \bar{GP}_t/1000$$

   El modelo de regresión es el mismo pero con las variables medidas en millones:

$$V_t^* = \alpha_0 + \alpha_1\, GP_t^* + u_t \qquad t = 1, 2, \ldots, 54$$

   Estimando los coeficientes por MCO se obtiene:

$$\hat{\alpha}_1 = \frac{\sum(V_t^* - \bar{V}^*)(GP_t^* - \bar{GP}^*)}{\sum(GP_t^* - \bar{GP}^*)^2} = \frac{\sum(V_t - \bar{V})(GP_t - \bar{GP})/(1000 \times 1000)}{\sum(GP_t - \bar{GP})^2/(1000 \times 1000)} = \hat{\beta}_1$$

$$\hat{\alpha}_0 = \bar{V}^* - \hat{\alpha}_1 \bar{GP}^* = \bar{V}/1000 - \hat{\beta}_1\, \bar{GP}/1000 = \hat{\beta}_0/1000$$

   De forma análoga los residuos al depender del tamaño de las variables quedarían divididos por (1000×1000).

   La recta de regresión muestral con las variables medidas en millones es:

$$\hat{V}_t^* = \underset{(0,143)}{0,48883} + \underset{(0,13)}{1,43}\; GP_t^* \qquad R^2 = 0,71 \quad SCR = 6,134381 \qquad t = 1907, \ldots, 1960$$

5. Estima por intervalo (probabilidad del 95 %) los coeficientes del modelo.

   Intervalo de confianza del 95 %: $[\hat{\beta}_i \pm t_{0,025}(T-(k+1)) \times \hat{\sigma}_{\hat{\beta}_i}]$.

   $\beta_0 \quad [488,83 \pm 2,007 \times 127,44] = [233,11 \quad 744,58]$

   $\beta_1 \quad [1,43 \pm 2,007 \times 0,13] = [1,18 \quad 1,69]$

6. Contrasta la significación estadística de la variable gastos en publicidad.

$$H_0: \quad \beta_1 = 0$$
$$H_a: \quad \beta_1 \neq 0$$

$$\mathrm{t} = \frac{\hat{\beta}_1 - 0}{\hat{\sigma}_{\hat{\beta}_1}} \sim t(T-(k+1))$$

$$\mathrm{t} = \frac{1,43-0}{0,13} = 11,31 \qquad |t| = |11,31| > t_{0,025}(54-2) = 2,007$$

   Se rechaza la hipótesis nula con un nivel de significación del 5 %. Por lo tanto, se rechaza la hipótesis de que los gastos en publicidad no influyen en las ventas.

7. ¿Son consistentes los resultados de los apartados 5 y 6?

   Sí son consistentes ya que en el apartado 5 se ha comprobado que el valor $\beta_1 = 0$ no pertenece al intervalo de confianza del 95 % de probabilidad y en el apartado 6 se ha rechazado la hipótesis nula de que $\beta_1 = 0$ con un nivel de significación $\alpha = 5\,\%$.

8. Contrasta la hipótesis de que el efecto de la publicidad en las ventas es mayor que la unidad.

$$H_0: \quad \beta_1 = 1$$
$$H_a: \quad \beta_1 > 1$$

$$\mathrm{t} = \frac{\hat{\beta}_1 - 1}{\hat{\sigma}_{\hat{\beta}_1}} \sim t(T-(k+1))$$

$$\mathrm{t} = \frac{1,43-1}{0,13} = 3,31 \qquad t = 3,31 > t_{0,05}(54-2) = 1,67$$

   Se rechaza la hipótesis nula con un nivel de significación del 5 %. Por lo tanto, el efecto de la publicidad en las ventas es mayor que la unidad con un nivel de significación del 5 %.

9. Contrasta la hipótesis de que el efecto de la publicidad en las ventas es menor que cinco.

$$H_0: \quad \beta_1 = 5$$
$$H_a: \quad \beta_1 < 5$$

$$\mathrm{t} = \frac{\hat{\beta}_1 - 5}{\hat{\sigma}_{\hat{\beta}_1}} \sim t(T-(k+1))$$

$$\mathrm{t} = \frac{1,43-5}{0,13} = -27,46 \qquad t = -27,26 < -t_{0,05}(54-2) = -1,67$$

   Se rechaza la hipótesis nula con un nivel de significación del 5 %. Por lo tanto, el efecto de la publicidad en las ventas es menor que cinco con un nivel de significación del 5 %.

10. Predice por punto las ventas del producto para el año 1961 si se ha presupuestado un gasto en publicidad de 600.

    La predicción por punto se obtiene utilizando la recta de regresión muestral:

    $$\hat{V}_{1961} = 488,83 + 1,43 \times 600 = 1349,59 \text{ miles de dólares.}$$

11. Proporciona un intervalo de predicción con una confianza del 95 % para las ventas del año 1962 si se ha presupuestado un gasto en publicidad de 575. Interpreta el resultado.

    Intervalo de predicción del 95 % de probabilidad: $[\hat{V}_{1962} \pm t_{0,025}(T-(k+1)) \times \hat{\sigma}_{\hat{u}_{1962}}]$.

    donde: $\hat{\sigma}^2_{\hat{u}_{1962}} = \hat{\sigma}^2_u \; [1 + X'_{1962}(X'X)^{-1}X_{1962}]$

    $$\hat{V}_{1962} = 488,83 + 1,43 \times 575 = 1313,72 \text{ miles de dólares.}$$

    Intervalo de predicción (95 %): $[1313,72 \pm 2,007 \times 349,5565] = [612,16 \quad 2015,28]$

PARTE 2 (4 puntos)

Supongamos que sobre las ventas del producto de la empresa Lydia E. Pinkham además de los gastos en publicidad puede influir la renta disponible. Si suponemos que la relación entre las variables es lineal:

1. Especifica el modelo de regresión lineal apropiado y estímalo por MCO con los datos del fichero `lidia.gdt`.

   MRLG: $V_t = \beta_0 + \beta_1 \, GP_t + \beta_2 \, R_t + u_t \qquad t = 1907, \ldots, 1960$

   RRM: $\hat{V}_t = \underset{(144,78)}{393,76} + \underset{(0,13)}{1,45} \; GP_t + + \underset{(0,50)}{0,68} \; R_t \qquad R^2 = 0,72 \quad SCR = 5923189$

2. Interpreta los coeficientes del modelo.

   - $\hat{\beta}_0$: ventas estimadas en miles de dólares cuando los gastos en publicidad y la renta son cero.
   - $\hat{\beta}_1$: incremento estimado en el volumen de ventas (en miles de dólares) cuando los gastos en publicidad aumentan en mil dólares y la renta se mantiene constante.
   - $\hat{\beta}_1$: incremento estimado en el volumen de ventas (en miles de dólares) cuando la renta aumenta en 1 unidad y los gastos en publicidad se mantienen constantes.

3. Calcula una medida de bondad del ajuste e interprétala.

   El coeficiente de determinación $R^2 = 0,72$ mide la bondad de ajuste del modelo: el 72 % de la variabilidad de las ventas en la muestra queda explicada por la variabilidad en los gastos en publicidad y en la renta en términos lineales.

4. Contrasta la significatividad individual y conjunta de las variables explicativas del modelo.

$$\text{Publicidad} \qquad \begin{array}{ll} H_0: & \beta_1 = 0 \\ H_a: & \beta_1 \neq 0 \end{array}$$

$$\text{t} = \frac{1,45 - 0}{0,13} = 11,47 \qquad |t| = |11,47| > t_{0,025}(54-3) = 2,008$$

Se rechaza la hipótesis nula con un nivel de significación del 5 %.

$$\text{Renta} \qquad \begin{array}{ll} H_0: & \beta_2 = 0 \\ H_a: & \beta_2 \neq 0 \end{array}$$

$$\text{t} = \frac{0,68 - 0}{0,50} = 1,35 \qquad |t| = |1,35| < t_{0,025}(54-3) = 2,008$$

No se rechaza la hipótesis nula con un nivel de significación del 5 %.

Por lo tanto, la variable publicidad sí es individualmente significativa mientras que la variable renta no lo es.

$$\text{Publicidad y renta} \qquad \begin{array}{ll} H_0: & \beta_1 = \beta_2 = 0 \\ H_a: & \beta_1 \neq 0 \,\text{y/o}\, \beta_2 \neq 0 \end{array}$$

El estadístico de contraste es:

$$F = \frac{R^2/2}{(1-R^2)/(54-3)} \sim \mathcal{F}(2,51)$$

$$F = 65,85 > \mathcal{F}_{0,05}(2,51) = 3,18$$

Se rechaza la hipótesis nula con un nivel de significación del 5 %. Las variables renta y publicidad son conjuntamente significativas.

PARTE 3 (8 puntos)

Por otro lado, se plantea la hipótesis de que es razonable pensar que el efecto de la publicidad en las ventas no es constante independientemente del nivel de gasto en publicidad, sino que existen rendimientos decrecientes, es decir, a mayor nivel de gasto en publicidad menor es el efecto marginal de un incremento unitario de la publicidad sobre las ventas. Para ello, se plantea el siguiente modelo:

$$(3) \qquad \ln V_t = \beta_0 + \beta_1 \ln GP_t + \beta_2 \ln R_t + u_t$$

1. ¿Cumple el modelo (3) los supuestos del modelo de regresión lineal general? Estima el modelo propuesto y escribe la recta de regresión muestral.

   El modelo (3) cumple el supuesto de linealidad del modelo, ya que este supuesto implica linealidad en los coeficientes y no en las variables.

$$\widehat{\ln V_t} = \underset{(0,40)}{2,05} + \underset{(0,06)}{0,75} \ln GP_t + \underset{(0,03)}{0,08} \ln R_t \qquad R^2 = 0,7896 \quad SCR = 1,296$$

2. Interpreta los coeficientes estimados.

   Como el modelo es doble-logarítmico los coeficientes de pendiente (los que acompañan a las variables explicativas) tienen la interpretación de elasticidades.

   - $\hat{\beta}_0$: ventas estimadas en logaritmos cuando los gastos en publicidad y la renta, ambos en logaritmos, son cero.
   - $\hat{\beta}_1$: variación porcentual estimada en el volumen de ventas cuando los gastos en publicidad aumentan en un 1 % y la renta se mantiene constante.
   - $\hat{\beta}_2$: variación porcentual estimada en el volumen de ventas cuando la renta aumenta en un 1 % y los gastos en publicidad se mantienen constantes.

3. Calcula una medida de bondad del ajuste e interprétala. ¿Tiene la misma interpretación el coeficiente de determinación que en el modelo de la Parte 2?

   El coeficiente de determinación $R^2 = 0,7896$ mide la bondad de ajuste del modelo: el 72 % de la variabilidad de los logaritmos de las ventas en la muestra queda explicada por la variabilidad en el logaritmos de los gastos en publicidad y en el logaritmo de la renta.

   Se puede observar que la interpretación de este coeficiente de determinación no es la misma que en el modelo de la de la Parte 2. El coeficiente de determinación mide la bondad de ajuste del modelo en términos lineales. Así en el modelo (2) mide el porcentaje de la variación muestral de las ventas explicada por su relación lineal con la publicidad y la renta, mientras que en el modelo (3) mide la relación entre las variables en logaritmos.

4. Contrasta la significatividad individual de las variables explicativas.

$$\text{Publicidad} \qquad \begin{aligned} H_0 &: \quad \beta_1 = 0 \\ H_a &: \quad \beta_1 \neq 0 \end{aligned}$$

$$\text{t} = \frac{0,75 - 0}{0,06} = 12,97 \qquad\qquad |t| = |12,97| > t_{0,025}(54-3) = 2,008$$

   Se rechaza la hipótesis nula con un nivel de significación del 5 %.

$$\text{Renta} \qquad \begin{aligned} H_0 &: \quad \beta_2 = 0 \\ H_a &: \quad \beta_2 \neq 0 \end{aligned}$$

$$\text{t} = \frac{0,08 - 0}{0,03} = 2,6 \qquad\qquad |t| = |2,6| > t_{0,025}(54-3) = 2,008$$

   Se rechaza la hipótesis nula con un nivel de significación del 5 %.

   Por lo tanto, las variables publicidad y renta son individualmente significativas.

5. Contrasta si la elasticidad gastos de publicidad-ventas es mayor a la unidad.

$$\begin{aligned} H_0 &: \quad \beta_1 = 1 \\ H_a &: \quad \beta_1 > 1 \end{aligned}$$

$$\text{t} = \frac{\hat{\beta}_1 - 1}{\hat{\sigma}_{\hat{\beta}_1}} \sim t(T-(k+1))$$

$$\text{t} = \frac{0,75 - 1}{0,06} = -4,17 \qquad\qquad t = -4,17 < t_{0,05}(54-2) = 1,67$$

No se rechaza la hipótesis nula con un nivel de significación del 5 %. Por lo tanto, el efecto de la publicidad en las ventas no es mayor que la unidad con un nivel de significación del 5 %.

6. ¿Las variables explicativas son conjuntamente significativas? Realiza el contraste oportuno ($\alpha = 5\,\%$).

$$H_0: \quad \beta_1 = \beta_2 = 0$$
$$H_a: \quad \beta_1 \neq 0 \,\text{y/o}\, \beta_2 \neq 0$$

El estadístico de contraste es:

$$F = \frac{R^2/2}{(1-R^2)/(54-3)} \sim \mathcal{F}(2,51) \qquad\qquad F = 93,86 > \mathcal{F}_{0,05}(2,51) = 3,18$$

Se rechaza la hipótesis nula con un nivel de significación del 5 %. Las variables renta y publicidad son conjuntamente significativas.

7. Dados los resultados obtenidos en la estimación del modelo (3), se piensa que la elasticidad renta-ventas es igual que la elasticidad gastos de publicidad-ventas. Contrasta la hipótesis propuesta.

$$H_0: \quad \beta_1 = \beta_2$$
$$H_a: \quad \beta_1 \neq \beta_2$$

El estadístico de contraste es:

$$F = \frac{(SCR_R - SCR_{NR})/1}{SCR_{NR}/(54-3)} \sim \mathcal{F}(1, 54-3)$$

donde

- Modelo No restringido: Modelo (3).
- Modelo Restringido:

$$\ln V_t = \beta_0 + \beta_1 \ln GP_t + \beta_1 \ln R_t + u_t = \beta_0 + \beta_1 (\ln GP_t + \ln R_t) + u_t$$

Una vez estimados ambos modelos con la muestra disponible, se obtienen las correspondientes sumas de cuadrados de los residuos y se puede computar el estadístico F muestral.

$$F = 93,54 > \mathcal{F}_{0,05}(1,51) = 4,03$$

Se rechaza la hipótesis nula con un nivel de significación del 5 % y concluimos que las dos elasticidades no son iguales.

8. Teniendo en cuenta los resultados que has obtenido en el contraste del apartado anterior, ¿qué modelo estimarías para explicar las ventas el modelo (3) o el modelo restringido que has propuesto y que incluye la hipótesis planteada? Razona tu respuesta.

El estimador de Mínimos Cuadrados Restringidos (es decir, incluyendo la restricción en el modelo) siempre tiene menor varianza que el de MCO, pero es sesgado si la restricción es falsa. Como en el contraste del apartado anterior se ha rechazado la hipótesis de que las elasticidades son iguales, los estimadores del modelo restringido son sesgados y el modelo adecuado sería el modelo (3).