Tarea 2: Modelos Estadísticos I.

Rojas Gutiérrez Rodolfo Emmanuel

24 de febrero de 2021

1. Ejercicios.

Ejercicio 1. (a) Pruebe que las estimaciones de mínimo cuadráticas $a_1, \ldots, a_m, b$ de los parámetros $\alpha_1, \ldots, \alpha_m, \beta$ de la familia de rectas

$$E[Y_i] = \alpha_i + \beta X_i,\ i = 1, \ldots, m,$$

están dados por

$$b = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n_i}(X_{ij} - \overline{X_i})(Y_{ij} - \overline{Y_i})}{\sum_{i=1}^{m}\sum_{j=1}^{n_i}(X_{ij} - \overline{X_i})^2}, \tag{1}$$

$$a_i = \overline{Y}_i - b\overline{X}_i, , \tag{2}$$

donde $(X_{i1}, Y_{i1}), \ldots, (X_{in_i}, Y_{in_i})$ denotan los valores observados de (X_i, Y_i) relacionados con la i-ésima recta, $i = 1, \ldots, m$.

(b) Pruebe que la suma de cuadrados residual está dada por

$$S^2 = \sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij} - \overline{Y_i}) - b^2\sum_{i=1}^{m}\sum_{j=1}^{n_i}(X_{ij} - \overline{X_i})^2, \tag{3}$$

con $\sum_{i=1}^{m} n_i - m - 1$ grados de libertad.

Solución. **(a)** Sea $(X_{11}, Y_{11}), \ldots, (X_{1n_1}, Y_{1n_1}), \ldots, (X_{m1}, Y_{m1}), \ldots, (X_{mn_m}, Y_{mn_m})$ una muestra observada, donde $(X_{i1}, Y_{i1}), \ldots, (X_{in_i}, Y_{in_i})$ denotan los valores observados de (X_i, Y_i) relacionados con la i-ésima recta, $i = 1, \ldots, m$. Defina para $i \in \{1, \ldots, m\}$

$$\overline{Y}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} Y_{ij}, \quad \overline{X}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} X_{ij},$$

entonces se cumple para cada $i \in \{1, \ldots, m\}$

$$\sum_{j=1}^{n_i}(Y_{ij} - \overline{Y}_i)^2 = \sum_{j=1}^{n_i} Y_{ij}^2 - n_i\overline{Y}_i^2, \tag{4}$$

$$\sum_{j=1}^{n_i}(X_{ij} - \overline{X}_i)^2 = \sum_{j=1}^{n_i} X_{ij}^2 - n_i\overline{X}_i^2, \tag{5}$$

$$\sum_{j=1}^{n_i}(X_{ij} - \overline{X}_i)(Y_{ij} - \overline{Y}_i) = \sum_{j=1}^{n_i} X_{ij}Y_{ij} - n_i\overline{X}_i\overline{Y}_i, \tag{6}$$

lo anterior servirá a lo largo de la prueba. Como se desea obtener las estimaciones por mínimos cuadrados de los parámetros $\alpha_1, \ldots, \alpha_m, \beta$ entonces se intentará minimizar la función $f : \mathbb{R}^{m+1} \to \mathbb{R}^+$ con regla de correspondencia

$$f(a_1, \ldots, a_m, b) = \sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij} - a_i - bX_{ij})^2, \tag{7}$$

para ello se buscará el punto en el que $\nabla f = 0$. Con esto en mente observe que para cada $r \in \{1, \ldots, m\}$ se satisface que

$$\frac{\partial f}{\partial a_r} = -2\sum_{j=1}^{n_r}(Y_{rj} - a_r - bX_{rj}), \tag{8}$$

por otro lado, note que

$$\frac{\partial f}{\partial b} = -2\sum_{i=1}^{m}\sum_{j=1}^{n_i} X_{ij}\left[Y_{ij} - a_i - bX_{ij}\right], \tag{9}$$

igualando las expresiones en (8) a cero obtenemos que

$$0 = \frac{\partial f}{\partial a_r} \iff \sum_{j=1}^{n_r}(Y_{rj} - a_r - bX_{rj}) = 0,$$

$$-n_r a_r + n_r\overline{Y}_r - bn_r\overline{X}_r = 0 \iff a_r = \overline{Y}_r - b\overline{X}_r, \tag{10}$$

por otra parte, igualando la expresión (9) a cero y multiplicando por $-\frac{1}{2}$ la desigualdad resultante se obtiene

$$0 = \sum_{i=1}^{m}\sum_{j=1}^{n_i} X_{ij}\left[Y_{ij} - a_i - bX_{ij}\right] = \sum_{i=1}^{m}\left[\sum_{j=1}^{n_i} X_{ij}Y_{ij} - a_i n_i\overline{X}_i - b\sum_{j=1}^{n_i} X_{ij}^2\right]$$

Usando (10) para sustituir el termino a_i se obtiene que.

$$= \sum_{i=1}^{m}\left[\sum_{j=1}^{n_i} X_{ij}Y_{ij} - n_i\left(\overline{Y}_i - b\overline{X}_i\right)\overline{X}_i - b\sum_{j=1}^{n_i} X_{ij}^2\right]$$

$$= \sum_{i=1}^{m}\left[\sum_{j=1}^{n_i} X_{ij}Y_{ij} - n_i\overline{X}_i\overline{Y}_i - b\left(\sum_{j=1}^{n_i} X_{ij}^2 - n_i\overline{X}_i^2\right)\right] = (*)$$

$$(*) = \underbrace{\sum_{i=1}^{m}\left[\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_i)(Y_{ij}-\overline{Y}_i) - b\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_i)^2\right]}_{\text{Por (5) y (6).}} \tag{11}$$

de este modo

$$\sum_{i=1}^{m}\left[\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_i)(Y_{ij}-\overline{Y}_i) - b\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_i)^2\right] = 0,$$

así pues

$$b = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_i)(Y_{ij}-\overline{Y}_i)}{\sum_{i=1}^{m}\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_i)^2}, \tag{12}$$

y de (10) se tenía que

$$a_r = \overline{Y}_r - b\overline{X}_r,\ r \in \{1,\ldots,m\}.$$

por ende, la solución a $\nabla f = 0$ es precisamente la expresada anteriormente, y lo anterior son los estimadores por mínimos cuadrados de los interceptos y el parámetro de pendiente.

(b) Observe que los residuales en este caso están dados por

$$e_{ij} = Y_{ij} - a_i - bX_{ij}, \quad \text{con } i = \{1,\ldots,m\} \text{ y } j = \{1,\ldots,n_i\},$$

de este modo la suma de cuadrados residuales S^2 esta dada por

$$S^2 = \sum_{i=1}^{m}\sum_{j=1}^{n_i} e_{ij}^2 = \sum_{i=1}^{m}\sum_{j=1}^{n_i}\left(Y_{ij} - a_i - bX_{ij}\right)^2$$

Usando (10) para sustituir el termino a_i se obtiene que.

$$= \sum_{i=1}^{m}\sum_{j=1}^{n_i}\left(Y_{ij} - \overline{Y}_i - b(X_{ij}-\overline{X}_i)\right)^2. \tag{13}$$

Por otro lado observe que de (12) se sigue que

$$b\sum_{i=1}^{m}\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_i)^2 = \sum_{i=1}^{m}\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_i)(Y_{ij}-\overline{Y}_i), \tag{14}$$

de este modo

$$\begin{aligned}\sum_{i=1}^{m}\sum_{j=1}^{n_i}\left(Y_{ij} - \overline{Y}_i - b(X_{ij}-\overline{X}_i)\right)^2 &= \sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij}-\overline{Y}_i)^2 - 2b\sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij}-\overline{Y}_i)(X_{ij}-\overline{X}_i) + b^2\sum_{i=1}^{m}\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_i)^2 \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij}-\overline{Y}_i)^2 \underbrace{-2b^2\sum_{i=1}^{m}\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_i)^2}_{\text{Por (14).}} + b^2\sum_{i=1}^{m}\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_i)^2 \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij}-\overline{Y}_i)^2 - b^2\sum_{i=1}^{m}\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_i)^2, \end{aligned} \tag{15}$$

de (13) y (15) se sigue

$$S^2 = \sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij} - \overline{Y}_i)^2 - b^2\sum_{i=1}^{m}\sum_{j=1}^{n_i}(X_{ij} - \overline{X}_i)^2,$$

con $\sum_{i=1}^{m} n_i - m - 1$ grados de libertad, el número de datos observados menos el número de parámetros a estimar. ■

Ejercicio 2. Los datos de la siguiente tabla muestran la relación entre la frecuencia cardiaca en reposo (Y) y el peso corporal en kilogramos X

(a) Grafique estos datos. ¿Parece que hay una relación lineal entre el peso corporal y la frecuencia cardiaca en reposo?

(b) Calcule las estimaciones $\hat{\beta}_0$ y $\hat{\beta}_1$ y escriba la ecuación de regresión para estos datos. Grafique la línea de regresión sobre la gráfica de dispersión del inciso **(a)** interprete los coeficientes de regresión estimados.

(c) Ahora examine el punto $(67, 40)$. Si dicho punto fuera eliminado del conjunto de datos, ¿qué cambios resultarian en las estimaciones del intercepto y la pendiente?

(d) Obtenga la estimación puntual de la respuesta media cuando $x = 88$. Obtenga un intervalo de $95\,\%$ de confianza para la respuesta media $X = 88kg$. Interprete este intervalo.

(e) Pronostique la frecuencia cardiaca para una persona cuyo peso corporal sea de $88kg$ utilizando una predicción puntual y una por intervalo del $95\,\%$ de confianza. Compare estas predicciones con la estimación del inciso **(d)**.

(f) Sin hacer cálculos, ¿para que valor medido de X la correspondiente $\hat{Y}$ tendría la menor varianza? ¿Por qué?

Solución. **(a)** La gráfica solicitada se presenta en la figura 1 en la misma se puede decir que no se nota a primer instancia una relación lineal entre las observaciones.

(b) Con los datos proporcionados el coeficiente de pendiente puede calcularse como

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{6} X_i Y_i - 6\overline{XY}}{\sum_{i=1}^{6} X_i^2 - 6\overline{x}^2} \approx 0.595, \tag{16}$$

por otro lado se sabe que

$$\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1\overline{X} \approx 4.799, \tag{17}$$

por lo que la ecuación de la recta de regresión esta dada por

$$\widehat{E}[Y_i|X_i] = 4.799 + 0.595X_i, \tag{18}$$

la gráfica de la recta de regresión ajustada sobre los datos puede consultarse en la figura 2. Por otro lado podemos interpretar el coeficiente de pendiente $\hat{\beta}_1$ como el cambio estimado en la media

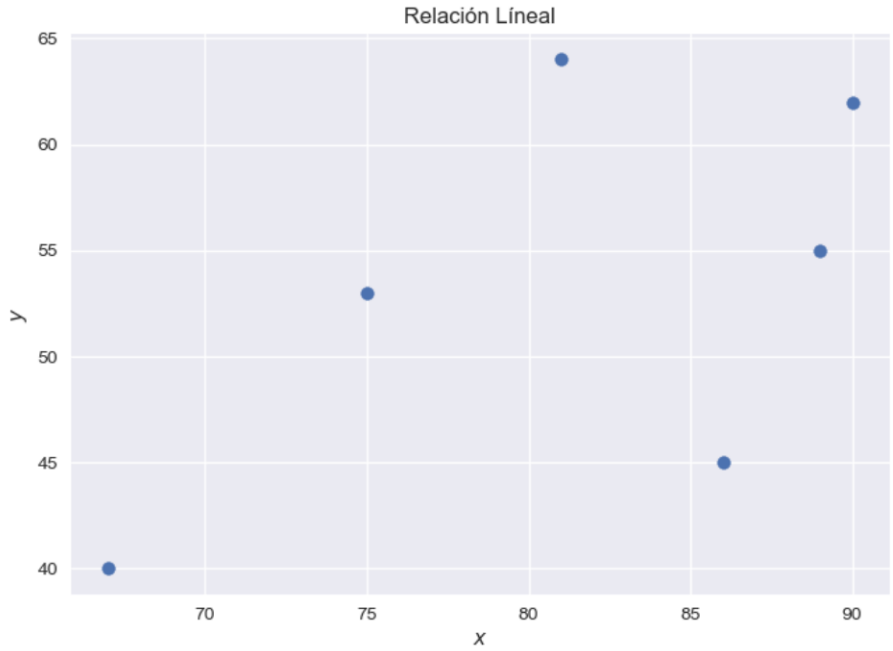

Figura 1: Relación Lineal.

de la frecuencia cardiaca dada una variación de un kilogramo en el peso, por ejemplo si se aumenta un kilo en el peso entonces la frecuencia cardiaca se elevara en 0.595 latidos por minuto. Por otra parte, el intercepto puede ser interpretado como la estimación de la media de la frecuencia cardiaca para un nivel de peso de 0 kilogramos, pese a ello esta interpretación es un poco inútil dentro del contexto del problema.

(c) Sí se eliminara el punto $(67, 40)$ se tendrían las siguientes estimaciones para β_1 y β_0

$$\hat{\beta}_1^* = \frac{\sum_{i=1}^{6} X_i Y_i - 67 \cdot 40 - (5)\left(\frac{\sum X_i - 67}{5}\frac{\sum Y_i - 40}{5}\right)}{\sum_{i=1}^{6} X_i^2 - 67^2 - 5\left(\frac{\sum X_i - 67}{5}\right)^2} \approx 0.079, \tag{19}$$

por otro lado, se sabe que

$$\hat{\beta}_0^* = \frac{\sum Y_i - 40}{5} - \hat{\beta}_1 \frac{\sum X_i - 67}{5} \approx 49.164, \tag{20}$$

por lo que, la ecuación de la recta de regresión en este caso esta dada por

$$\widehat{E}[Y_i | X_i] = 49.164 + 0.079 X_i, \tag{21}$$

comparando los estimadores calculados en (16) y (17) con los que fueron calculados en (19) y (20), se puede notar como el intercepto creció bastante y el término de pendiente se hizo mucho más cercano a cero, esto se debe a que si observa con detenimiento la figura 2 es fácil percatarse que el

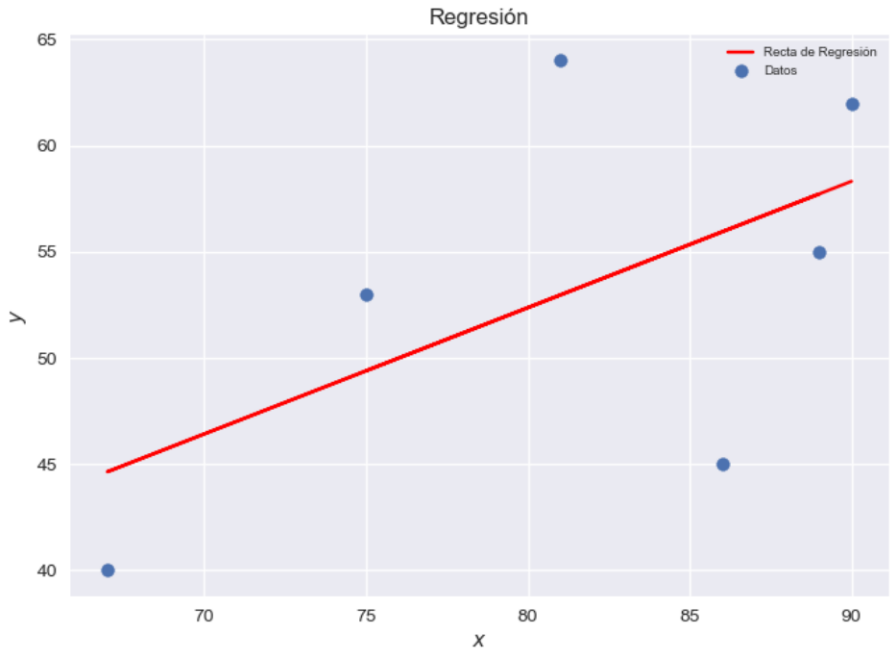

Figura 2: Recta de Regresión Sobre el Conjunto de Datos.

punto removido es aquel que queda en la esquina inferior izquierda, el mismo parece ser el culpable de que la recta de regresión quede inclinada de esa manera y por ende la pendiente sea mas pronunciada en ese caso.

(d) La estimación puntual para la media de la frecuencia cardiaca dado un valor de $X = 88$ está dada por

$$\widehat{E}(Y|X = 88) = \hat{\beta}_0 + \hat{\beta}_1 \cdot 88 \approx 57.131. \tag{22}$$

Con $\hat{\beta}_1, \hat{\beta}_0$ como en (16) y (17). Mientras que para realizar la estimación por intervalos primero se calculara la cantidad

$$s = \sqrt{\frac{SS(Res)}{n-2}} = \sqrt{\frac{\sum_{i=1}^{6} Y_i^2 - 6\overline{Y}^2 - \hat{\beta}_1^2\left[\sum_{i=1}^{6} X_i^2 - 6\overline{X}^2\right]}{4}} \approx 8.615. \tag{23}$$

Sea $\alpha = 0.05$ y $t = t_{4,\frac{\alpha}{2}}$ el cuantil $\frac{\alpha}{2}$ de una distribución t con 4 grados de libertad entonces el intervalo de confianza del 95 % para la media de la frecuencia cardiaca, dado un nivel del peso de 88 kilogramos, esta dado por

$$\left[\widehat{E}(Y|X = 88) + t \cdot s \cdot \sqrt{\frac{1}{6} + \frac{(88 - \overline{X})^2}{\sum_{i=1} X_i^2 - 6\overline{X}^2}}, \widehat{E}(Y|X = 88) - t \cdot s \cdot \sqrt{\frac{1}{6} + \frac{(88 - \overline{X})^2}{\sum_{i=1} X_i^2 - 6\overline{X}^2}}\right]$$
$$= [44.533, 69.730], \tag{24}$$

recuerde que el intervalo de confianza de 95 % de confianza es un intervalo aleatorio, donde el nivel de confianza hace referencia a que el mismo contiene en un 95 % de las veces al verdadero valor de la media de la respuesta dado un valor de la variable independiente de $X = 88kg$, por ende con los datos con los que se ajusto el modelo se tendrá con un 95 % de confianza que la media de la frecuencia cardiaca de una persona cuyo peso es de $88kg$, se encuentra entre 44.533 y 69.730 latidos por minuto.

(e) La estimación puntual para la frecuencia cardiaca dado un valor de $X = 88$, también conocida como pronóstico o predicción, está nuevamente dada por

$$\hat{Y}_{pred} = \widehat{E}(Y|X = 88) = 88 = \hat{\beta}_0 + \hat{\beta}_1 \cdot 88 \approx 57.131.$$

Con $\hat{\beta}_1, \hat{\beta}_0$ como en (16) y (17). Mientras que para calcular los intervalos de predicción se utilizará el estadístico s (23), y nuevamente sea $\alpha = 0.05$ y $t = t_{4,\frac{\alpha}{2}}$ el cuantil $\frac{\alpha}{2}$ de una distribución t con 4 grados de libertad entonces el intervalo de predicción con confianza del 95 % para la frecuencia cardiaca, dado un nivel del peso de 88 kilogramos esta dado por

$$\left[\widehat{E}(Y|X = 88) + t \cdot s \cdot \sqrt{1 + \frac{1}{6} + \frac{(88 - \overline{X})^2}{\sum_{i=1} X_i^2 - 6\overline{X}^2}}, \widehat{E}(Y|X = 88) - t \cdot s \cdot \sqrt{1 + \frac{1}{6} + \frac{(88 - \overline{X})^2}{\sum_{i=1} X_i^2 - 6\overline{X}^2}}\right]$$
$$= [30.096, 84.166], \tag{25}$$

Observe que las estimaciones puntuales son idénticas, mientras que el intervalo de predicción es ligeramente más ancho que el intervalo de confianza, esto se debe a que el valor de la predicción siempre posee un variabilidad mayor a la de su valor promedio.

(f) Viendo a $\hat{Y}$ como estimador de la media de la frecuencia cardíaca para cierto nivel de peso X, observe que la varianza de Y está dada por

$$V(\hat{Y}) = \sqrt{\frac{1}{n} + \frac{(X - \overline{X})}{\sum_{i=1}^{n}(X_i - \overline{X})^2}} \geq \sqrt{\frac{1}{n}},$$

donde el valor a la derecha se alcanza sí y solo si $X = \overline{X}$. De igual modo si se toma a $\hat{Y}$ como predicción de la frecuencia cardiaca para cierto nivel de peso X, se tiene que su varianza está dada por

$$V(\hat{Y}) = \sqrt{1 + \frac{1}{n} + \frac{(X - \overline{X})}{\sum_{i=1}^{n}(X_i - \overline{X})^2}} \geq \sqrt{1 + \frac{1}{n}},$$

donde el valor a la derecha se alcanza sí y solo si $X = \overline{X}$. Por lo que, el valor que minimiza la varianza de $\hat{Y}$ es justamente la media de las X_i.

Por último en la figura 3, se deja una gráfica de la recta de regresión lineal ajustada sobre los datos con los intervalos de predicción y confianza correspondientes.

■

Ejercicio 3. Use los datos del ejercicio 2 y calcule el valor de $\hat{Y}_i$ para cada valor de X. Calcule las correlaciones entre

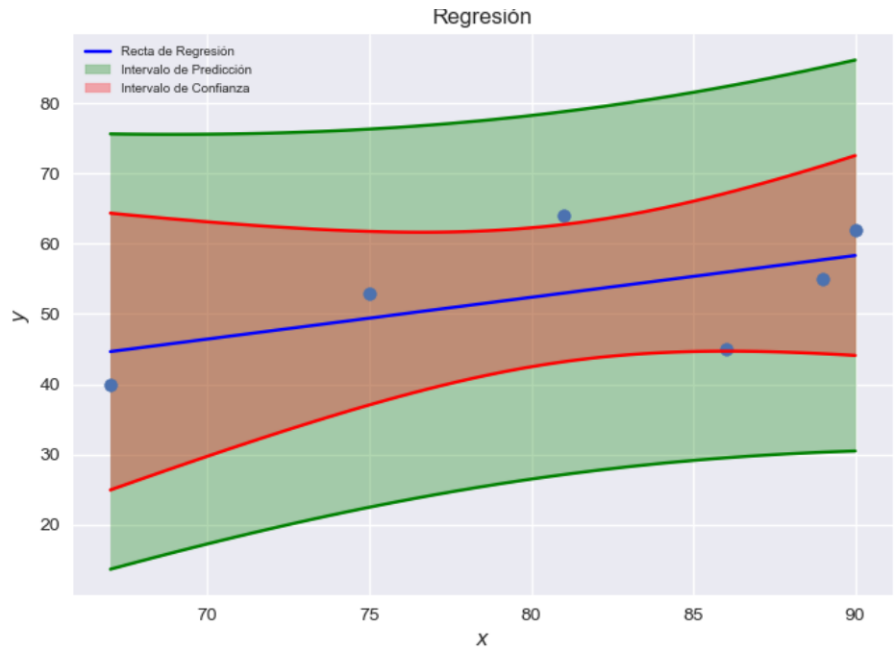

Figura 3: Gráfica con intervalos.

(a) X_i y Y_i.

(b) Y_i y $\hat{Y}_i$.

(c) X_i y $\hat{Y}_i$.

Compare estas correlaciones entre ellas y con el coeficiente de determinación R^2. ¿Puede probar algebraícamente las relaciones que detectó.

Solución. **(a)** La correlación entre X_i e Y_i es aproximadamente igual a 0.569

(b) La correlación entre Y_i y $\hat{Y}_i$ es aproximadamente igual a 0.569

(c) La correlación entre X_i y $\hat{Y}_i$ es 1.0

.

Para poder realizar las comparaciones solicitadas se calculo el coeficiente de determinación

$$R^2 = \frac{SS(Reg)}{SS(Tot)} \approx 0.323. \tag{26}$$

Los cálculos de todas estas cantidades pueden ser consultadas en el script adjunto a esta tarea. Usando los incisos **(a)**, **(c)** y (26) es fácil notar que el valor de R^2 es el cuadrado del coeficiente de correlación entre X_i e Y_i, al igual que el cuadrado del coeficiente de correlación entre Y_i y $\hat{Y}_i$. Por

otro lado, no se encuentra una relación clara entre el coeficiente de correlación entre X_i e $\hat{Y}_i$ y los demás valores solicitados. Por otra parte, se demostrará que en el caso de la regresión lineal simple en la cual se considera un intercepto se cumple que

$$R^2 = corr^2(X_i, Y_i) = corr^2(Y_i, \hat{Y}_i).$$

Para ello recuerde que

$$\begin{aligned} SS(Reg) &= \hat{\beta}_1^2 \sum_{i=1}^{n} (X_i - \overline{X})^2. \\ \hat{\beta}_1 &= \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n}(X_i - \overline{X})^2}, \end{aligned} \tag{27}$$

y

$$\begin{aligned} R^2 = \frac{SS(Reg)}{\sum_{i=1}^{n}(Y_i - \overline{Y})^2} &= \underbrace{\frac{\hat{\beta}_1^2 \sum_{i=1}^{n}(X_i - \overline{X})^2}{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}}_{\text{Por (27).}} \\ &= \left[\frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n}(X_i - \overline{X})^2}\right]^2 \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{\sum_{i=1}^{n}(Y_i - \overline{Y})^2} \\ &= \left[\frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2 \sum_{i=1}^{n}(Y_i - \overline{Y})^2}}\right]^2 = \left(corr(X, Y)\right)^2 \end{aligned} \tag{28}$$

por ende se cumple que

$$R^2 = \left(corr(X, Y)\right)^2, \tag{29}$$

por otro lado, observe que

$$\left(corr(Y, \hat{Y})\right)^2 = \left[\frac{\sum_{i=1}(\hat{Y}_i - \overline{\hat{Y}})(Y_i - \overline{Y}_i)}{\sqrt{\sum_{i=1}(\hat{Y}_i - \overline{\hat{Y}})^2 \sum_{i=1}^{n}(Y_i - \overline{Y}_i)^2}}\right]^2, \tag{30}$$

por último note que

$$\begin{aligned} \hat{Y}_i - \overline{\hat{Y}} &= \hat{\beta}_0 + \hat{\beta}_1 X_i - \frac{1}{n}\sum_{j=1}^{n}(\hat{\beta}_0 + \hat{\beta}_1 X_j) \\ &= \hat{\beta}_0 + \hat{\beta}_1 X_i - \hat{\beta}_0 - \hat{\beta}_1 \overline{X} \\ &= \hat{\beta}_1 (X_i - \overline{X}), \end{aligned} \tag{31}$$

de (30) y (31) se sigue que

$$\left(corr(Y, \hat{Y})\right)^2 = \left[\frac{\sum_{i=1}\hat{\beta}_1(X_i - \overline{X})(Y_i - \overline{Y}_i)}{\sqrt{\sum_{i=1}\hat{\beta}_1^2(X_i - \overline{X})^2 \sum_{i=1}^{n}(Y_i - \overline{Y}_i)^2}}\right]^2 = (*) \tag{32}$$

$$(*) = \left[\frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2 \sum_{i=1}^{n}(Y_i - \overline{Y})^2}} \right]^2 = (corr(X,Y))^2 , \tag{33}$$

de (29) y (33) se sigue que

$$(corr(Y, \hat{Y}))^2 = R^2,$$

lo que prueba las dos relaciones establecidas. ■

Ejercicio 4. Pruebe las siguientes relaciones

$$SS(Modelo) = n\overline{Y}^2 + \hat{\beta}_1^2 \sum_{i=1}^{n}(X_i - \overline{X})^2, \tag{34}$$

$$\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 = \sum_{i=1}^{n}(Y_i - \overline{Y})^2 - \hat{\beta}_1^2 \sum_{i=1}^{n}(X_i - \overline{X})^2. \tag{35}$$

Solución. Para probar (34) observe que

$$\begin{aligned} SS(Modelo) &= \sum_{i=1}^{n} \hat{Y}_i^2 = \sum_{i=1}^{n}(\hat{\beta}_0 + \hat{\beta}_1 X_i)^2 = \sum_{i=1}^{n}(\overline{Y} - \hat{\beta}_1 \overline{X} + \hat{\beta}_1 X_i)^2 = \sum_{i=1}^{n}(\overline{Y} - \hat{\beta}_1 (X_i - \overline{X}))^2 \\ &= \sum_{i=1}^{n} \overline{Y}^2 - 2\hat{\beta}_1 \overline{Y} \underbrace{\sum_{i=1}^{n}(X_i - \overline{X})}_{0} + \hat{\beta}_1^2 \sum_{i=1}^{n}(X_i - \overline{X})^2 \qquad (36) \\ &= n\overline{Y}^2 + \hat{\beta}_1^2 \sum_{i=1}^{n}(X_i - \overline{X})^2, \qquad (37) \end{aligned}$$

en (36) observe que $\sum_{i=1}^{n}(X_i - \overline{X}) = \sum_{i=1}^{n}(X_i) - n\overline{X} = 0$. De (37) se sigue que

$$SS(Modelo) = n\overline{Y}^2 + \hat{\beta}_1^2 \sum_{i=1}^{n}(X_i - \overline{X})^2. \tag{38}$$

Lo que prueba (34). Por otro lado para (35) observe que

$$\sum_{i=1}^{n}(Y_i - \overline{Y})^2 = \sum_{i=1}^{n} Y_i^2 - n\overline{Y}^2 = SS(Tot_{Nc}) - n\overline{Y}^2, \tag{39}$$

y de lo visto en clase se sabe que

$$SS(Tot_{Nc}) - n\overline{Y}^2 = \left[SS(Modelo) - n\overline{Y}^2 \right] + SS(Res), \tag{40}$$

en este punto se podría pensar se este abusando un poco, debido a que la igualdad anterior se da gracias a que $\sum_{i=1}^{n} \hat{Y}_i e_i = 0$, pero este hecho se probará en el siguiente ejercicio por lo que se usará

la igualdad vista en clase sin reparó alguno. De (39) y (40) se sigue que

$$\begin{aligned}\sum_{i=1}^{n}(Y_i-\overline{Y})^2 &= \left[SS(Modelo)-n\overline{Y}^2\right]+SS(Res)\\ &= \underbrace{\hat{\beta}_1^2\sum_{i=1}^{n}(X_i-\overline{X})^2}_{\text{Por (38).}}+SS(Res)\\ &= \hat{\beta}_1^2\sum_{i=1}^{n}(X_i-\overline{X})^2+\sum_{i=1}^{n}(Y_i-\hat{Y}_i)^2, \end{aligned} \tag{41}$$

De (41) se sigue

$$\sum_{i=1}^{n}(Y_i-\hat{Y}_i)^2=\sum_{i=1}^{n}(Y_i-\overline{Y})^2-\hat{\beta}_1^2\sum_{i=1}^{n}(X_i-\overline{X})^2.$$

Lo que prueba (35) y concluye el ejercicio. ■

Ejercicio 5. Pruebe algebraícamente que cuando la ecuación de regresión lineal simple contiene intercepto, se tiene que $\sum_{i=1}^{n} e_i = 0$. Muestre que esto no es así si la regresión no contempla el intercepto.

Solución. Suponga que se tiene un conjunto de observaciones $\{(X_i, Y_i)\}_{i=1}^{n}$, donde se conoce que no hay errores de medición en las mediciones de X y la incertidumbre recae sobre las observaciones de Y. Entonces para un modelo de regresión lineal simple que considera intercepto se tiene que los estimadores para sus coeficiente están dados por

$$\begin{aligned}\hat{\beta}_1 &= \frac{\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})}{\sum_{i=1}^{n}(X_i-\overline{X})^2},\\ \hat{\beta}_2 &= \overline{Y}-\hat{\beta}_1\overline{X}, \end{aligned} \tag{42}$$

por lo que

$$\begin{aligned}\sum_{i=1}^{n} e_i = \sum_{i=1}^{n}(Y_i-\hat{Y}_i) &= \sum_{i=1}^{n}\left(Y_i-\hat{\beta}_0-\hat{\beta}_1X_i\right)\\ &= \underbrace{\sum_{i=1}^{n}\left(Y_i-\overline{Y}+\hat{\beta}_1\overline{X}-\hat{\beta}_1X_i\right)}_{\text{Por (42)}}\\ &= n\overline{Y}-n\overline{Y}-\hat{\beta}_1\underbrace{\sum_{i=1}^{n}(X_i-\overline{X})}_{0} \qquad (43)\\ &= 0. \qquad (44)\end{aligned}$$

Donde en (43) observe que $\sum_{i=1}^{n}(X_i - \overline{X}) = \sum_{i=1}^{n}(X_i) - n\overline{X} = 0$. Mientras que para el modelo de regresión lineal simple que no considera intercepto, se satisface que el estimador de mínimos cuadrados para el coeficiente de pendiente esta dado por

$$\hat{\beta}_1^* = \frac{\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1} X_i^2}, \tag{45}$$

por lo que

$$\begin{aligned}\sum_{i=1}^{n} e_i &= \sum_{i=1}^{n}(Y_i - \hat{Y}_i) = \sum_{i=1}^{n}\left(Y_i - \hat{\beta}_1^* X_i\right) \\ &= \underbrace{\sum_{i=1}^{n}\left(Y_i - X_i \frac{\sum_{j=1}^{n} X_j Y_j}{\sum_{j=1} X_j^2}\right)}_{\text{Por (45)}} \\ &= \sum_{i=1}^{n} Y_i - \frac{\sum_{i=1}^{n} X_i \sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} X_i^2}, \end{aligned} \tag{46}$$

De (46) se sigue

$$\sum_{i=1}^{n} e_i = 0 \text{ sí y solo sí } \sum_{i=1}^{n} Y_i = \frac{\sum_{i=1}^{n} X_i \sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} X_i^2},$$

por lo que en general no es cierta esta igualdad a cero. Un contraejemplo sencillo podría darse considerando los siguientes datos $(1,0),(0,1)$ con esta información

$$\sum_{i=1}^{2} Y_i = 1,$$

pero

$$\frac{\sum_{i=1}^{2} X_i \sum_{i=1}^{2} X_i Y_i}{\sum_{i=1}^{2} X_i^2} = 0,$$

■

Ejercicio 6. Use las ecuaciones del modelo de regresión lineal simple para probar que

(a) $\sum_{i=1}^{n} X_i Y_i = \sum X_i \hat{Y}_i$.
(b) $\sum_{i=1}^{n} X_i e_i = 0$
(c) $\sum_{i=1}^{n} \hat{Y}_i e_i = 0$

Solución. Suponga que se tiene un conjunto de observaciones $\{(X_i, Y_i)\}_{i=1}^{n}$, donde se conoce que no hay errores de medición en las mediciones de X y la incertidumbre recae sobre las observaciones de Y. Entonces para un modelo de regresión lineal simple que considera intercepto se tiene que los estimadores para sus coeficiente están dados por

$$\begin{aligned}\hat{\beta}_1 &= \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n}(X_i - \overline{X})^2}, \\ \hat{\beta}_0 &= \overline{Y} - \hat{\beta}_1 \overline{X}, \end{aligned} \tag{47}$$

(a) Observe que

$$\begin{aligned}\sum_{i=1}^{n} X_i\hat{Y}_i &= \sum_{i=1}^{n} X_i\left(\hat{\beta}_0+\hat{\beta}_1X_i\right) = \hat{\beta}_0\sum_{i=1}^{n}X_i+\hat{\beta}_1\sum_{i=1}^{n}X_i^2\\ &= \underbrace{n\overline{X}(\overline{Y}-\hat{\beta}_1\overline{X})}_{\text{Por (47)}}+\hat{\beta}_1\sum_{i=1}^{n}X_i^2 = n\overline{XY}-\hat{\beta}_1n\overline{X}^2+\hat{\beta}_1\sum_{i=1}^{n}X_i^2\\ &= n\overline{XY}+\hat{\beta}_1\sum_{i=1}^{n}(X_i-\overline{X})^2 = n\overline{XY}+\underbrace{\frac{\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})}{\sum_{i=1}^{n}(X_i-\overline{X})^2}\sum_{i=1}^{n}(X_i-\overline{X})^2}_{\text{Nuevamente por (47).}}\\ &= n\overline{XY}+\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y}) = n\overline{XY}+\sum_{i=1}^{n}X_iY_i-n\overline{XY} = \sum_{i=1}^{n}X_iY_i, \qquad (48)\end{aligned}$$

de (48) se sigue

$$\sum_{i=1}^{n} X_i\hat{Y}_i = \sum_{i=1}^{n} X_iY_i.$$

(b) Este inciso se sigue directamente del inciso **(a)** ya que

$$\sum_{i=1}^{n} X_ie_i = \sum_{i=1}^{n} X_i\left(Y_i-\hat{Y}_i\right) = \underbrace{\sum_{i=1}^{n}X_iY_i-\sum_{i=1}^{n}X_i\hat{Y}_i}_{\text{Por el inciso (a).}} = 0, \qquad (49)$$

de (49) se sigue que

$$\sum_{i=1}^{n} X_ie_i = 0.$$

(c) Este inciso se sigue directamente del inciso **(b)** ya que

$$\sum_{i=1}^{n}\hat{Y}_ie_i = \sum_{i=1}^{n}(\hat{\beta}_0+\hat{\beta}_1X_i)e_i = \hat{\beta}_0\sum_{i=1}^{n}e_i+\hat{\beta}_1\sum_{i=1}^{n}X_ie_i = 0, \qquad (50)$$

ya que $\sum_{i=1}^{n} e_i = 0$ por ser un modelo que considera intercepto, y por lo probado en el ejercicio 5 y $\sum_{i=1}^{n} X_ie_i = 0$ por el inciso **(b)**. De (50) se sigue que

$$\sum_{i=1}^{n}\hat{Y}_ie_i = 0,$$

lo que concluye el inciso **(c)** y el ejercicio. ■

Ejercicio 7. Obtenga las ecuaciones normales y las estimaciones mínimo cuadráticas para el modelo

$$Y_i = \mu+\beta_1x_i+\varepsilon_i, \qquad (51)$$

donde $x_i = X_i - \overline{X}$. Compare estos resultados con los que se obtienen en el modelo de regresión lineal usual.

Solución. Suponga que se tiene un conjunto de observaciones $\{(X_i, Y_i)\}_{i=1}^n$, donde se conoce que no hay errores de medición en las mediciones de X y la incertidumbre recae sobre las observaciones de Y. Para ajustar el modelo centrado se define $x_i = X_i - \overline{X}$ y se buscará encontrar las estimaciones por mínimos cuadrados minimizando la función $f : \mathbb{R}^2 \to \mathbb{R}$ con regla de correspondencia

$$f(\mu, \beta_1) = \sum_{j=1}^{n} \left(Y_i - \mu - \beta_1 x_i\right)^2, \tag{52}$$

como la anterior función es diferenciable bastará con resolver la ecuación $\nabla f = 0$. Con esto en mente, observe que

$$\frac{\partial f}{\partial \mu} = -2 \sum_{j=1}^{n} \left(Y_i - \mu - \beta_1 x_i\right), \tag{53}$$

por otro lado se cumple que

$$\frac{\partial f}{\partial \beta_1} = -2 \sum_{j=1}^{n} x_i \left(Y_i - \mu - \beta_1 x_i\right). \tag{54}$$

Igualando a cero la expresión (53) y multiplicando por $-\frac{1}{2}$ se tiene que

$$0 = \sum_{j=1}^{n} (Y_i - \mu - \beta_1 x_i) = n\overline{Y} - n\mu - \underbrace{\beta_1 n \overline{x}}_{0} \tag{55}$$

$$= n\overline{Y} - n\mu \tag{56}$$

donde en (55) se tiene que $\overline{x} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}) = 0$. Por último de (56) se sigue que

$$\hat{\mu} = \overline{Y}, \tag{57}$$

igualando ahora a cero la expresión en (54), multiplicando por $-\frac{1}{2}$ y sustituyendo el valor encontrado para μ (57) se sigue que

$$\begin{aligned} 0 &= \sum_{i=1}^{n} x_i \left(Y_i - \hat{\mu} - \beta_1 x_i\right) \\ &= \sum_{i=1}^{n} x_i \left(Y_i - \overline{Y} - \beta_1 x_i\right) \\ &= \sum_{i=1}^{n} x_i (Y_i - \overline{Y}) - \beta_1 \sum_{i=1}^{n} x_i^2, \end{aligned}$$

de lo anterior se sigue que

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} x_i (Y_i - \overline{Y})}{\sum_{i=1}^{n} x_i^2} = \frac{\sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n} (X_i - \overline{X})^2},$$

de esto último y de (57) se tiene que la suma de los errores cuadráticos se minimiza en $(\hat{\mu}, \hat{\beta}_1)$ con

$$\begin{aligned} \hat{\mu} &= \overline{Y}, \\ \hat{\beta}_1 &= \frac{\sum_{i=1}^{n} x_i (Y_i - \overline{Y})}{\sum_{i=1}^{n} x_i^2} = \frac{\sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n} (X_i - \overline{X})^2}, \end{aligned} \tag{58}$$

por lo que, en (58) se exponen los estimadores por mínimos cuadrados solicitados. Comparando con los de la regresión lineal con modelo no centrado, cuyos estimadores se presentan a continuación

$$\begin{aligned}\hat{\beta}_0^* &= \overline{Y} - \hat{\beta}_1^* \overline{X},\\ \hat{\beta}_1^* &= \frac{\sum_{i=1}^n (X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^n (X_i - \overline{X})^2},\end{aligned} \tag{59}$$

La única diferencia que se presenta entre los estimadores en (58) y en (59), es que el estimador para el intercepto no depende del parámetro de pendiente, esto se debe a que la media de los datos centrados es cero, note que esto coincide con el aplicar el modelo no centrado a datos cuyo promedio es cero. Por último, para evitar romper con la continuidad de la solución se exponen las ecuaciones normales para este caso, igualando las dos derivadas parciales encontradas en (53) y (54) a cero y realizando un par de despejes

$$\begin{aligned} n\hat{\mu} + \hat{\beta}_1 \sum_{i=1}^n x_i &= \sum_{i=1}^n Y_i,\\ \hat{\mu} \sum_{i=1}^n x_i + \hat{\beta}_1 \sum_{i=1}^n x_i^2 &= \sum_{i=1}^n x_i Y_i,\end{aligned}$$

o en términos de las X

$$\begin{aligned} n\hat{\mu} + \hat{\beta}_1 \sum_{i=1}^n (X_i - \overline{X}) &= \sum_{i=1}^n Y_i,\\ \hat{\mu} \sum_{i=1}^n (X_i - \overline{X}) + \hat{\beta}_1 \sum_{i=1}^n (X_i - \overline{X})^2 &= \sum_{i=1}^n (X_i - \overline{X}) Y_i.\end{aligned}$$

■

Ejercicio 8. Recalcule la ecuación de regresión y el análisis de varianza para el ejemplo de los datos de ozono contra rendimiento de las plantas de soya usando el modelo centrado

$$Y_i = \mu + \beta_1 x_i + \varepsilon_i,$$

donde $x_i = X_i - \overline{X}$. Compare y comente los resultados que se obtienen con ambos modelos

Ozono (ppm) X	Y Rendimiento (gm/plt)
0.02	242
0.07	237
0.11	231
0.15	201

Cuadro 1: Datos Rendimiento de Soya.

Solución. Del ejercicio anterior se sabe que las estimaciones para el intercepto y el coeficiente de pendiente para este modelo están dadas por

$$\hat{\mu} = \overline{Y},$$
$$\hat{\beta}_1^* = \frac{\sum_{i=1}^n x_i(Y_i - \overline{Y})}{\sum_{i=1}^n x_i^2} = \frac{\sum_{i=1}^n (X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^n (X_i - \overline{X})^2},$$

de este modo utilizando los datos en la tabla 1 se obtiene que

$$\hat{\mu} = 227.75 \text{ y } \hat{\beta}_1^* = -293.531, \tag{60}$$

de este modo se tiene que la recta de regresión en este caso esta dada por

$$\widehat{E}(Y_i|X_i) = 227.75 - 293.531 x_i = 227.75 - 293.531(X_i - \overline{X}), \tag{61}$$

observando que $\overline{X} = 0.0875$ y usando (61) se tiene que

$$\widehat{E}(Y_i|X_i) = 253.434 - 293.531 X_i, \tag{62}$$

que es básicamente la recta de regresión obtenida en clase, por lo que se esta haciendo la misma estimación solo que parámetrizada de otra manera. Por otra parte, el análisis de varianza se detalla en la tabla 2, los cálculos para el mismo se pueden observar en el script adjunto a esta tarea. Como puede observarse el análisis de varianzas es prácticamente idéntico al obtenido en clase, lo cual no debería resultar sorprendente por el hecho de que se esta utilizando un modelo reparametrizado.

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrado medio
Total	3	1014.75	
Debida a Reg.	1	799.138	799.138
Residual	2	215.612	107.806

Cuadro 2: Análisis de Varianzas Rendimiento de Soya.

■

Ejercicio 9. Pruebe que

$$t = \frac{\hat{\beta}_1}{\frac{s}{\sqrt{\sum_{i=1}^n (X_i - \overline{X})^2}}},$$

es igual a $\frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$ donde

$$s^2 = \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n-2} \text{ y } r = \frac{\sum_{i=1}^n (X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^n (X_i - \overline{X})^2 \sum_{i=1}^n (Y_i - \overline{Y})^2}}$$

Solución. Observe que

$$s^2 = \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n-2} = \underbrace{\frac{\sum_{i=1}^n (Y_i - \overline{Y})^2 - \hat{\beta}_1^2 \sum_{i=1}^n (X_i - \overline{X})^2}{n-2}}_{\text{Por el ejercicio 4.}} \tag{63}$$

así pues

$$
\begin{aligned}
t &= \frac{\hat{\beta}_1}{\frac{s}{\sqrt{\sum_{i=1}^n (X_i-\overline{X})^2}}} = \hat{\beta}_1 \underbrace{\frac{\sqrt{n-2}\sqrt{\sum_{i=1}^n (X_i-\overline{X})^2}}{\sqrt{\sum_{i=1}^n (Y_i-\overline{Y})^2 - \hat{\beta}_1^2 \sum_{i=1}^n (X_i-\overline{X})^2}}}_{\text{Por (63).}} \\
&= \hat{\beta}_1 \frac{\sqrt{\sum_{i=1}^n (X_i-\overline{X})^2}\sqrt{n-2}}{\sqrt{\sum_{i=1}^n (Y_i-\overline{Y})^2 - \left(\frac{\sum_{i=1}^n (X_i-\overline{X})(Y_i-\overline{Y})}{\sum_{i=1}^n (X_i-\overline{X})^2}\right)^2 \sum_{i=1}^n (X_i-\overline{X})^2}} \\
&= \frac{\sum_{i=1}^n (X_i-\overline{X})(Y_i-\overline{Y})}{\sum_{i=1}^n (X_i-\overline{X})^2} \frac{\sqrt{n-2}\sqrt{\sum_{i=1}^n (X_i-\overline{X})^2}}{\sqrt{\sum_{i=1}^n (Y_i-\overline{Y})^2}\sqrt{1-\left[\frac{\sum_{i=1}^n (X_i-\overline{X})(Y_i-\overline{Y})}{\sqrt{\sum_{i=1}^n (Y_i-\overline{Y})^2 \sum_{i=1}^n (X_i-\overline{X})^2}}\right]^2}} \\
&= \frac{\sum_{i=1}^n (X_i-\overline{X})(Y_i-\overline{Y})}{\sqrt{\sum_{i=1}^n (X_i-\overline{X})^2 \sum_{i=1}^n (Y_i-\overline{Y})^2}} \frac{\sqrt{n-2}}{\sqrt{1-\left[\frac{\sum_{i=1}^n (X_i-\overline{X})(Y_i-\overline{Y})}{\sqrt{\sum_{i=1}^n (Y_i-\overline{Y})^2 \sum_{i=1}^n (X_i-\overline{X})^2}}\right]^2}} \\
&= \frac{r\sqrt{n-2}}{\sqrt{1-r^2}},
\end{aligned}
$$

de dondes se sigue que

$$t = \frac{\sqrt{n-2}r}{1-r^2},$$

el resultado deseado. ■

Ejercicio 10. Considere el modelo de regresión lineal

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i,\ i = 1, 2, \ldots, n,$$

donde $\varepsilon_1, \ldots, \varepsilon_n$ son variables aleatorias independientes con con distribución normal con media cero y varianza común σ^2.

(a) Obtenga los estimadores de máxima verosimilitud de los parámetros del modelo β_0, β_1 y σ^2.

(b) Compare los estimadores de máxima verosimilitud y de mínimos cuadrados. Comente esta comparación.

Solución. **(a)** Se considerará el modelo de regresión lineal

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i,\ i = 1, 2, \ldots, n, \tag{64}$$

con términos de error normales, es decir, los términos de error $\varepsilon_1, \ldots, \varepsilon_n$ son variables aleatorias con distribución normal de media 0 y varianza común σ^2, de este hecho y de (65) se sigue que

$Y_i \sim N(\beta_0 + \beta_1 X_i, \sigma^2)$ para cada $i \in \{1, \ldots, n\}$ así pues la función de verosimilitud de la muestra aleatoria $Y = (Y_1, \ldots, Y_n)$ deberá ser proporcional a

$$\begin{aligned} f(Y) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2\sigma^2}(Y_i - \beta_0 - \beta_1 X_i)^2\right\} \\ &= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 X_i)^2\right\} \\ &\propto \sigma^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 X_i)^2\right\}, \end{aligned} \tag{65}$$

así pues, la función de verosimilitud puede escribirse como

$$L(\sigma^2, \beta_0, \beta_1 | Y) = \sigma^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 X_i)^2\right\},$$

por lo que la función de $log - verosimilitud$ queda dada por

$$l(\sigma^2, \beta_0, \beta_1 | Y) = -\frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 X_i)^2, \tag{66}$$

para encontrar el máximo verosímil se buscara resolver las ecuaciones, i.e

$$0 = \frac{\partial l}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 X_i)^2, \tag{67}$$

$$0 = \frac{\partial l}{\partial \beta_0} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 X_i), \tag{68}$$

$$0 = \frac{\partial l}{\partial \beta_1} = \frac{1}{\sigma^2}\sum_{i=1}^{n} X_i(Y_i - \beta_0 - \beta_1 X_i), \tag{69}$$

resolviendo (68) se obtiene que

$$\begin{aligned} &\frac{1}{\sigma^2}\sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 X_i) = 0 \\ &\qquad\qquad \Longleftrightarrow \\ &n\overline{Y} - n\beta_0 - n\beta_1\overline{X} = 0 \iff \beta_0 = \overline{Y} - \beta_1\overline{X}, \end{aligned} \tag{70}$$

luego sustituyendo (70) en (69)

$$\frac{1}{\sigma^2}\sum_{i=1}^{n} X_i(Y_i - \beta_0 - \beta_1 X_i) = 0$$

$$\Longleftrightarrow$$

$$\begin{aligned} 0 &= \sum_{i=1}^{n} X_i Y_i - \sum_{i=1}^{n}(\overline{Y} - \beta_1)X_i - \beta_1 \sum_{i=1}^{n} X_i^2 \\ &= \sum_{i=1}^{n} X_i Y_i - n\overline{XY} + n\beta_1 \overline{X}^2 - \beta_1 \sum_{i=1}^{n} X_i^2 \\ &= \sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y}) - \beta_1 \sum_{i=1}^{n}(X_i - \overline{X})^2, \end{aligned}$$

de lo que se sigue que

$$\beta_1 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n}(X_i - \overline{X})^2}.$$

De lo anterior y de (70) se tiene que los estimadores de máxima verosimilitud para β_0 y β_1 están dados por

$$\begin{aligned} \hat{\beta}_0 &= \overline{Y} - \hat{\beta}_1 \overline{X}, \\ \hat{\beta}_1 &= \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n}(X_i - \overline{X})^2}, \end{aligned} \tag{71}$$

Por último, sustituyendo las expresiones (71) en (67) se tiene que

$$\frac{n}{2\sigma^2} = \frac{1}{2\sigma^4}\sum_{i=1}^{n}(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)^2,$$

y despejando σ^2 se sigue que

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)^2}{n} = \frac{SS(Res)}{n}, \tag{72}$$

es el estimador de máxima verosimilitud de σ^2.

(b) En (71) se observa que los estimadores de máxima verosimilitud para β_0 y β_1, en el caso de términos de error normales *i.i.d*, son exactamente iguales a los estimadores por mínimos cuadrados en el caso en el que solo se sabe que los errores son *i.i.d* con media 0 y varianza σ^2, sin embargo el estimador para σ^2 difiere del conocido estimador $s^2 = \frac{SS(Res)}{n-2}$, para la varianza utilizado en regresión por una constante multiplicativa, ya que

$$\hat{\sigma}^2_{emv} = \frac{n-2}{n} s^2.$$

. ■