

Capítulo 6

Crítica y Diagnóstico del Modelo Lineal

6.1 Coeficiente de Determinación Múltiple

Otra herramienta que puede emplearse para estudiar la adecuación de un modelo es el *Coeficiente de Determinación Múltiple* R^2 , el cual se define como

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}$$

De manera informal, podemos pensar que R^2 representa el porcentaje de la variación de los datos que es explicado por el modelo. Claramente, $0 < R^2 \leq 1$, y mientras más cerca esté R^2 de uno, mayor será la variación explicada.

Sin embargo, nótese que cada vez que se introduce una nueva variable independiente en el modelo aumenta SSR , y por tanto aumenta R^2 ; es decir, R^2 puede acercarse todo lo que se desee a uno sólo incorporando suficientes variables al modelo. Por lo tanto, el coeficiente de determinación múltiple debe ser usado con sumo cuidado, y no debe emplearse como único criterio de comparación entre modelos, ya que no penaliza la complejidad del modelo.

Como ilustración de este procedimiento, con los datos del ejemplo 5.1 calculemos los valores de R^2 para los dos modelos correspondientes al ejemplo 5.1 que fueron comparados en la sección 5.2.

Para el modelo (1), $y_i = \beta_0 + \beta_2 x_{2i} + \varepsilon_i$, (el cual explica el consumo del producto en función del nivel educativo en la zona) obtenemos

$$R_1^2 = \frac{154.41}{1279.2} = 0.5897$$

Es decir, el modelo (1) explica alrededor del 60% de la variación de los datos.

Para el modelo (2), $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \varepsilon$ (es decir, el modelo que incluye adicionalmente el nivel de urbanización y el ingreso relativo), se obtiene como valor para el coeficiente de determinación

$$R_2^2 = \frac{1081.35}{1279.2} = 0.8453$$

Es decir, el modelo (2) explica aproximadamente el 85% de la variación de los datos.

Parece entonces haber una gran diferencia entre ambos modelos. Sin embargo, cabe preguntarse si ese aumento del R^2 es debido a una mejoría real en el modelo, o sólo al hecho de introducir una variable adicional.

Para penalizar la complejidad del modelo, definimos el *Coficiente de Determinación Múltiple Corregido*, R^{2*} , como:

$$R^{2*} = 1 - \frac{\frac{SSE}{n-(k+1)}}{\frac{SST}{n-1}} = 1 - \frac{MSE}{MST}$$

Nótese que, a diferencia de R^2 , modelos que produzcan el mismo SSE pueden generar diferentes valores de R^{2*} , dependiendo de su complejidad.

Para los modelos anteriores, $R_1^{2*} = 0.5311$ y $R_2^{2*} = 0.7525$. En ambos casos, el coeficiente de determinación corregido da valores inferiores a los arrojados por el coeficiente de determinación múltiple. En particular, para el modelo (2) obtenemos una disminución apreciable. Sin embargo, aún se trata de comparar un 75% de variación explicada para el modelo (2) contra un 53% de variación explicada para el modelo (1). Esto puede llevarnos a pensar que, contrariamente a lo obtenido usando la prueba F, un modelo para explicar el consumo del producto debe incluir el nivel de urbanización o el ingreso relativo, o ambas.

6.2 Crítica del Modelo

Toda la inferencia que hemos desarrollado para modelos lineales descansa en cuatro suposiciones básicas:

- La relación entre las variables es lineal.
- Los errores siguen una distribución normal.
- Las varianzas de los errores son iguales (es decir, los errores son *homocedásticos*).
- Los errores son independientes.

Claramente, es necesario tener alguna manera de determinar si estas hipótesis se cumplen, al menos en forma aproximada. Para ello, resulta razonable usar los residuos

$$\mathbf{e} = \mathbf{Y} - \hat{\mathbf{Y}} = \mathbf{Y} - X\hat{\boldsymbol{\beta}},$$

pues éstos resumiran toda la información disponible sobre los errores.

En la sección 4.3 se probó que si el modelo $\mathbf{Y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ es correcto, los residuos pueden ser escritos como $\mathbf{e} = (I - V)\boldsymbol{\varepsilon}$, donde $V = X(X'X)^{-1}X'$ es la matrix de proyección sobre el espacio generado por las columnas de X . Es decir, si el modelo es correcto, los residuos son combinación lineal de los errores, y ésto implica que los residuos siguen una distribución Normal con media

$$\begin{aligned} E(\mathbf{e}) &= E[(I - V)\boldsymbol{\varepsilon}] \\ &= (I - V)E(\boldsymbol{\varepsilon}) \\ &= (I - V)\mathbf{0} \\ &= \mathbf{0} \end{aligned}$$

y varianza

$$\begin{aligned} Var(\mathbf{e}) &= E[(\mathbf{e} - E(\mathbf{e}))(\mathbf{e} - E(\mathbf{e}))'] \\ &= E[(\mathbf{e} - \mathbf{0})(\mathbf{e} - \mathbf{0})'] \\ &= E[(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}')] \\ &= E[(I - V)\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'(I - V)] && (I - V) \text{ es simétrica} \\ &= (I - V)E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}')(I - V) \\ &= \sigma^2(I - V) && (I - V) \text{ es idempotente} \end{aligned}$$

Por tanto, $Var(e_i) = \sigma^2(1 - v_{ii})$, donde v_{ii} es el i -ésimo elemento de la diagonal de V .

El diagnóstico del modelo se basa en comparar el comportamiento de los residuos con el comportamiento de variables aleatorias normales. A efectos prácticos, suelen analizarse los *residuos estandarizados*

$$r_i = \frac{e_i}{S\sqrt{1 - v_{ii}}},$$

es decir, los residuos divididos entre un estimador de su desviación estandard.

El análisis de los residuos estandarizados se hace en forma gráfica. Los gráficos de residuos más usados son los siguientes:

1. Gráfico de residuos en papel normal.

Para construir un gráfico probabilístico normal, ordenamos los residuos en forma ascendente y graficamos el k -ésimo residuo vs. el punto $Z_i = \Phi^{-1}\left(\frac{k-1/2}{n}\right)$, donde Φ es la función de distribución Normal(0,1).

Si los datos provienen de una distribución normal, este gráfico debería lucir como una línea recta. Cuando se detectan desviaciones sistemáticas a partir de este patrón (por ejemplo, forma de S o curvas) obtenemos evidencia de que los datos no provienen de una distribución normal. En este caso, es necesario transformar la variable dependiente para obtener un comportamiento más semejante al de un conjunto de datos normales.

Este gráfico también es útil para identificar valores atípicos.

2. Gráfico de Residuos vs. Valores Estimados

Si el modelo que se emplea es correcto, los residuos deberían carecer de estructura y lucir simplemente como ruido, pues toda la señal contenida en los datos ha sido explicada por el modelo. Por lo tanto, un gráfico de residuos satisfactorio en esta situación no deberá mostrar ningún patrón evidente de comportamiento, como se indica en la figura 6.1

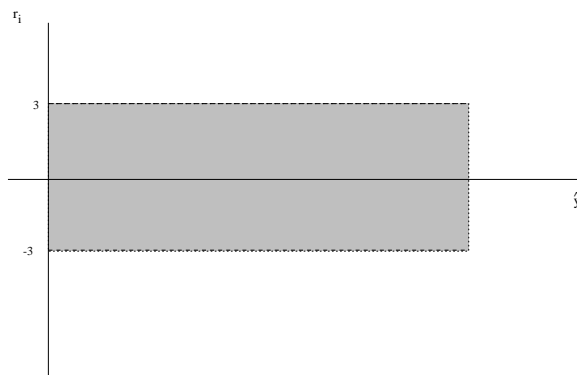


Figura 6.1: Forma de un gráfico satisfactorio de residuos.

Si usamos los residuos estandarizados r_i , es raro que los residuos se encuentren fuera del intervalo $(-3, 3)$, y debería haber aproximadamente el mismo número de residuos positivos y negativos.

Algunas formas típicas para gráficos de residuos insatisfactorios se muestran en la figura 6.2. En el caso (1), se evidencia que es necesario incluir otra variable explicativa, mientras que en el caso (2) es necesario incluir un término de orden superior (cuadrático, cúbico, etc.). Finalmente, el caso (3) se presenta cuando la varianza de los residuos no es constante, sino que depende del valor ajustado (es decir, los errores son *heterocedásticos*). Cuando esto sucede, nuevamente es necesario aplicar alguna transformación a la variable dependiente con el fin de estabilizar la varianza.

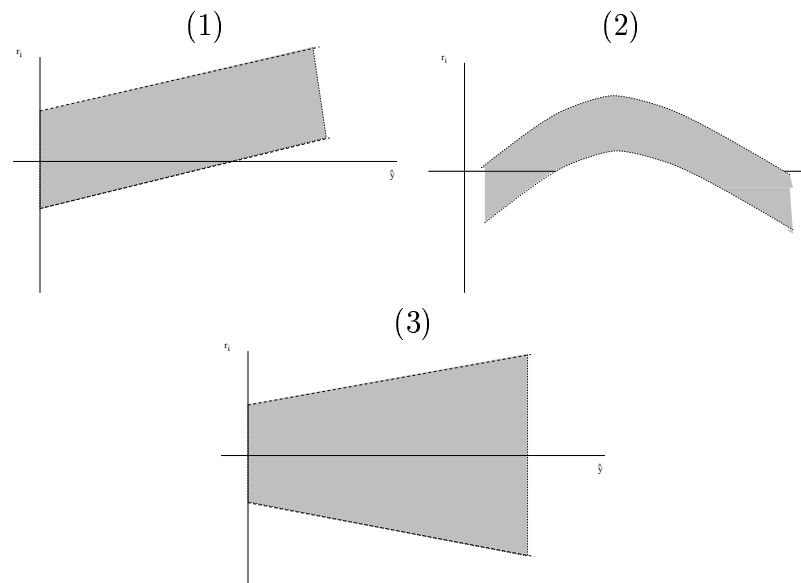


Figura 6.2: Formas típicas de gráficos de residuos insatisfactorios.

3. Gráfico de residuos vs. variables explicativas.

4. Gráfico de residuos vs. variables no incluidas en el modelo.

Cuando se presentan anomalías como las descritas en el punto 2., estos gráficos permiten determinar qué variable debe incluirse, o para cuál de ellas es necesario incluir un término de orden superior.

Ejemplo 6.1 Consideremos el siguiente conjunto de datos

X_1	10	20	30	12	14	16	18	22	24	26	28
X_2	2	1	4	3	2	2	4	5	1	3	5
Y	24.2	84	185.8	34.7	41	58	69	98.4	127	133	139.2

Usando un paquete estadístico se ajustó el modelo $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i$ y se obtuvieron los siguientes resultados

Coefficiente	Valor	Error Standard	Valor t
β_0	-59.58		
β_1	7.73	0.52	14.82
β_2	-1.64	2.40	-0.68

Fuente	Grados de Libertad	SS	MS	F
Modelo	2	25381.91	12690.95	128.04
Error	8	792.98	99.12	
TOTAL	10	26174.89		

Si comparamos el valor calculado de F con el valor tabulado $F_{2,8}^{0.025} = 6.06$ concluimos que el modelo ajustado es mejor que el modelo $y_i = \beta_0 + \varepsilon_i$. Los valores $R^2 = 0.9697$ y $R^{2*} = 0.9621$ indican, además, que el modelo explica aproximadamente el 97% de la variación de los datos.

Al comparar los valores t para los distintos coeficientes con el valor tabulado $t_8^{0.025} = 2.36$ observamos que el siguiente paso parecería ser eliminar x_2 del modelo. Sin embargo, antes de continuar analicemos los residuos obtenidos al ajustar este modelo, los cuales se presentan en la figura 6.3. En el gráfico de residuos vs valores ajustados se observa un patrón de curva; cuando graficamos los residuos con respecto a X_1 y X_2 , queda claro que ese patrón corresponde a la variable X_1 .

En vista de los resultados obtenidos, decidimos ajustar un modelo de la forma:

$$y_i = \beta_0 + \beta_1 x_{1i}^2 + \beta_2 x_{2i} + \varepsilon_i$$

El paquete estadístico nos da

Coefficiente	Valor	Error Standard	Valor t
β_0	12.19		
β_1	0.20	0.01	22.35
β_2	-2.70	1.62	-1.67

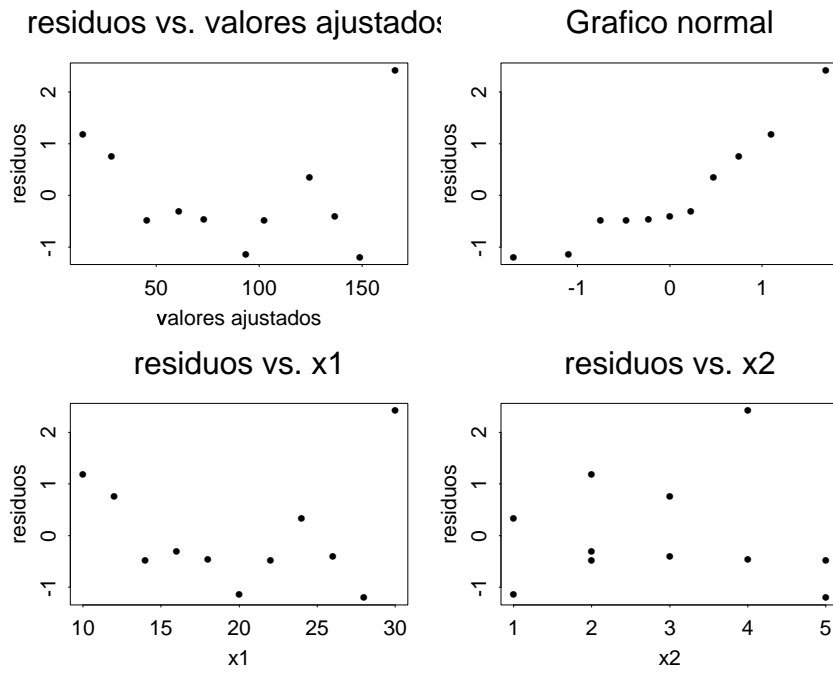


Figura 6.3: Gráficos de residuos para el ejemplo 6.1

Fuente	Grados de Libertad	SS	MS	F
Modelo	2	25819.32	12909.66	290.43
Error	8	355.57	44.45	
TOTAL	10	26174.89		

Nuevamente se obtiene un modelo mejor que el modelo $y = \text{constante} + \text{error}$. Incluso aumenta el R^2 (en este caso es 0.9864) para un modelo de la misma complejidad. Sin embargo, es necesario continuar el análisis, pues pareciera razonable eliminar X_2 del modelo en base al valor del t correspondiente.

Los gráficos de residuos se presentan en la figura 6.4. Allí puede verse que la tendencia en los residuos ha desaparecido, si bien el gráfico normal de residuos revela la presencia de un valor cuyo residuo se encuentra inusualmente alejado del resto. Aún así, es claro que este segundo modelo es preferible al primero.

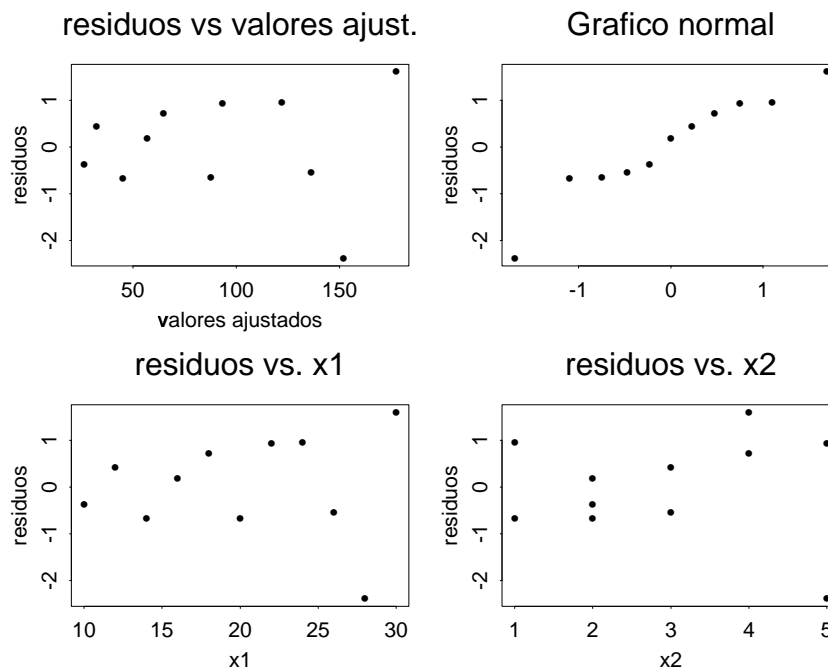


Figura 6.4: Gráficos de residuos para el segundo modelo, ejemplo 6.1

6.3 Selección del Mejor Conjunto de Variables Independientes

6.3.1 Covarianza y Correlación

Recordemos que la covarianza para dos variables aleatorias X y Y se define como

$$\text{Cov}(X, Y) = E[(X - E(X))(Y - E(Y))]$$

La covarianza es una medida de la relación lineal entre dos variables aleatorias, pero tiene el inconveniente de depender de la magnitud de los datos. Para eliminar dicha dependencia se define el *coeficiente de correlación* ρ como

$$\rho = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

Puede probarse que $|\rho| \leq 1$, de manera que mientras más cercana a 1 esté una correlación, mayor será la asociación *lineal* entre las variables en estudio. Una correlación $\rho > 0$ indica una relación lineal positiva, es decir, X crece cuando Y crece. Lo contrario sucede si $\rho < 0$: Y decrece cuando X crece.

El equivalente muestral de esta última medida es la *correlación muestral* r , la cual se define para n pares de observaciones $(X_1, Y_1), \dots, (X_n, Y_n)$ como

$$r(X, Y) = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

Nótese la semejanza con la definición de correlación para dos variables aleatorias (básicamente, se sustituyen esperanzas por promedios).

Al tratarse de una medida de asociación lineal, parece razonable incluir la correlación como parte del análisis de un modelo de regresión, e incorporar al modelo las variables explicativas que tengan la mayor correlación con Y . Sin embargo, este criterio puede presentar problemas cuando las variables independientes están cerca de ser linealmente dependientes; esta situación se conoce como *multicolinealidad*. En este caso, los estimadores de mínimos cuadrados no miden los efectos individuales, sino que reflejan un efecto parcial.

Nótese que, al ser la correlación muestral una medida de relación lineal entre dos vectores de datos, podemos usarla para medir relación lineal entre las variables independientes, aún cuando no se trate de variables aleatorias. Por lo tanto, hablaremos de la *matriz de correlación* para las variables del problema, y diremos que dos variables independientes están “correlacionadas” cuando exista relación lineal entre ellas.

En el Ejemplo 5.1 la matriz de correlación es la siguiente:

	X_1	X_2	X_3
X_1	1		
X_2	0.684	1	
X_3	-0.616	-0.172	1
Y	0.802	0.768	-0.629

Obsérvese que la variable independiente más correlacionada con Y es X_1 , pero ésta, a su vez, está muy correlacionada con las otras dos. Es decir, si incluimos X_1 , X_2 y X_3 en el modelo, estaremos introduciendo información en algún sentido “redundante”, por lo cual los coeficientes estimados no medirán los efectos individuales de cada variable. En efecto, veamos la matriz de correlación estimada para las componentes del vector $\hat{\beta}$ cuando se ajusta el modelo $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \varepsilon_i$

	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$
$\hat{\beta}_1$	1		
$\hat{\beta}_2$	-0.7442	1	
$\hat{\beta}_3$	0.6929	0.4326	1

Como era de esperarse, vemos que $\hat{\beta}_1$ está correlacionado con $\hat{\beta}_2$ y $\hat{\beta}_3$. Estas relaciones explican los fenómenos observados en este ejemplo al aplicar la prueba t para probar hipótesis sobre los parámetros del modelo.

6.4 Técnicas de regresión por pasos

Teniendo en cuenta el fenómeno de multicolinealidad, se han desarrollado diversos métodos automáticos para la selección del mejor conjunto de variables independientes. Muchos de estos métodos pueden clasificarse como métodos de *regresión por pasos*, pues implican la inclusión o eliminación de una variable independiente del modelo en cada paso del proceso.

A continuación estudiaremos dos de estas técnicas de regresión por pasos: la *eliminación regresiva* y la *inclusión progresiva*.

- **Eliminación Regresiva.**

La técnica de *eliminación regresiva* (Eliminación hacia atrás o “Backwards Regression”) funciona de la siguiente manera:

1. Se ajusta el modelo que contiene todas las posibles variables independientes.

2. Se toma el parámetro que tiene menor valor calculado para t , digamos β_i
3. Se prueba la hipótesis $H_0 : \beta_i = 0$. Si no se puede rechazar esta hipótesis, la variable X_i es eliminada del modelo, se ajusta un nuevo modelo y se regresa al paso 2.
4. Si la hipótesis $H_0 : \beta_i = 0$ es rechazada, el procedimiento termina.

Esta es la técnica que se usó al ajustar el modelo correspondiente al ejemplo 5.1

• **Inclusión Progresiva.**

El procedimiento de inclusión progresiva (Inclusión hacia adelante, “Stepwise Regression”) se basa en las correlaciones de las variables independientes X_1, X_2, \dots, X_k con la variable dependiente Y . Sin embargo, ya vimos que cuando las variables independientes están correlacionadas entre sí, este criterio puede llevar a conclusiones incorrectas.

Para superar este inconveniente, se procede de la siguiente manera: supongamos que se tiene el modelo $y_i = \beta_0 + \beta_1 x_{1i} + \varepsilon$. Se desea incluir una nueva variable explicativa entre X_2, X_3, \dots, X_k , pero se quiere “eliminar el efecto” de X_1 , es decir, se desea seleccionar una variable cuya correlación con Y no sea originada por una alta correlación con X_1 . Una manera de lograr este objetivo es calcular la correlación entre los residuos del modelo $y_i = \beta_0 + \beta_1 x_{1i} + \varepsilon$ y los residuos de cada uno de los modelos $x_{li} = \beta_0 + \beta_1 x_{1i} + \varepsilon$. A esta correlación la llamaremos *correlación parcial entre Y y X_l eliminando la influencia de X_1* , y la denotaremos como $r(Y - \hat{Y}(X_1), X_l - \hat{X}_l(X_1))$. Nótese además que puede eliminarse la influencia de tantas variables como se desee, simplemente incluyéndolas en el modelo.

El procedimiento de inclusión progresiva funciona entonces de la siguiente manera:

1. Se ajusta el modelo $y_i = \beta_0 + \beta_1 x_{1i} + \varepsilon$, donde X_1 es la variable independiente tal que $r(X_1, Y)$ sea máxima.
2. Se prueba la hipótesis $H_0 : \beta_1 = 0$. Si no se rechaza, el procedimiento termina.
3. Si se rechaza la hipótesis, se calculan las correlaciones parciales de Y con las variables restantes (eliminando la influencia de X_1), y se incluye en el modelo aquella con mayor correlación parcial.
4. Se prueba la hipótesis $H_0 : \beta_i = 0$ para todas las variables presentes en el modelo. Si no se rechaza esta hipótesis para la última variable incluida, el procedimiento termina.

5. Si se rechaza $H_0 : \beta_i = 0$ para la última variable incluida, pero no se rechaza para alguna otra variable, esta última se elimina del modelo, y nuevamente se trata de incluir otra variable.
6. El procedimiento se repite hasta que no se pueden incluir más variables en la ecuación.

Ejemplo 6.2 *Se desea estudiar el calor desprendido por un cierto tipo de cemento al endurecerse, en función de su composición. Las variables de interés en el estudio son:*

Y : Calor desprendido (cal/gr).

X_1 : Porcentaje de aluminato de calcio ($3CaO.Al_2O_3$).

X_2 : Porcentaje de silicato tetracálcico ($3CaO.SiO_2$).

X_3 : Porcentaje de ferrita aluminica tetracálcica ($4CaO.Al_2O_3.Fe_2O_3$).

X_4 : Porcentaje de silicato dicálcico ($2CaO.SiO_2$).

Al realizar un experimento, se obtienen los siguientes datos:

X_1	X_2	X_3	X_4	Y
7	26	6	60	78
1	29	15	52	74
11	56	8	20	104
11	31	8	47	87
7	52	6	33	95
11	55	9	22	109
3	71	17	6	102
1	31	22	44	72
21	47	4	26	115
2	54	18	22	93
1	40	23	34	83
11	66	9	12	113
10	68	8	12	109

Comencemos por explorar la matriz de correlación

	X_1	X_2	X_3	X_4
X_1	1			
X_2	0.229	1		
X_3	-0.824	-0.139	1	
X_4	-0.845	-0.973	0.3	1
Y	0.731	0.816	-0.535	-0.821

Observando estas correlaciones sería razonable esperar un modelo que no contenga simultáneamente los pares de variables $X_1 - X_3$, $X_1 - X_4$ o $X_2 - X_4$, pues éstas tienen correlaciones altas entre sí. En este caso, la variable más correlacionada con Y es X_4 .

Apliquemos el método de inclusión progresiva para el ajuste de este modelo.

- **PASO 1:** Incluir X_4 .

$$\text{Modelo: } y_i = \beta_0 + \beta_4 x_{4i} + \varepsilon_i$$

Para este modelo se obtiene $R^2 = 0.678$ y $F = 23.128$, con lo cual se rechaza la hipótesis $H_0 : \beta_4 = 0$.

Variable	Correlación Parcial	F-to-enter
X1	0.954	101.97
X2	0.132	0.178
X3	-0.896	40.752

Obsérvese la diferencia la correlación y la correlación parcial entre X_4 y X_2 .

El *F-to enter* es el valor de F que se obtiene al comparar el modelo actual con el modelo que también incluye la variable indicada. La información proporcionada por esta cantidad es equivalente a la correlación parcial.

- **PASO 2:** Incluir X_1 .

$$\text{Modelo: } y_i = \beta_0 + \beta_1 x_{1i} + \beta_4 x_{4i} + \varepsilon_i$$

Para este modelo se obtienen $R^2 = 0.971$ y $F = 168.690$. Nótese que hay un marcado incremento en el R^2 .

Parámetros	Estimados	F-to-remove
β_0	102.758	
β_1	1.43	101.966
β_4	-0.617	153.644

El *F-to-remove* corresponde a probar el modelo actual con el modelo sin la variable indicada (es, en realidad, el cuadrado del estadístico t correspondiente a dicha variable).

Obsérvese que la menor F es la correspondiente a la variable que acabamos de incluir, y su valor no permite eliminar dicha variable del modelo. Por tanto, tratamos de incluir una nueva variable.

Variable	Correlación Parcial	F-to-enter
X_2	0.584	4.775
X_3	-0.59	4.325

- **PASO 3:** Incluir X_2 .

Modelo: $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_4 x_{4i} + \varepsilon_i$

Para este modelo se obtienen $R^2 = 0.981$ (Obsérvese que casi no crece con respecto al modelo anterior) y $F = 156.502$. Nuevamente, el modelo ajustado es mejor que el modelo $y_i = \beta_0 + \varepsilon$.

Parámetros	Estimados	F-to-remove
β_0	71.121	
β_1	1.442	142.536
β_2	0.419	4.775
β_4	-0.237	1.757

El menor valor de F corresponde a β_4 ; al comparar este valor con el valor tabulado de $F_{1,9}^{0.05} = 4.67$, no rechazamos $H_0 : \beta_4 = 0$.

- **PASO 4:** Eliminar X_4 .

Modelo: $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i$

Para este modelo se obtienen $R^2 = 0.978$ (Obsérvese que casi no decrece con respecto al modelo anterior y es ligeramente mayor que el obtenido en el paso 2) y $F = 217.421$.

Parámetros	Estimados	F-to-remove
β_0	52.008	
β_1	1.458	136.98
β_2	0.665	194.558

El menor valor de F , correspondiente a β_1 , no permite eliminar X_1 del modelo.

Variables	Correlación Parcial	F-to-enter
X_3	0.387	1.585
X_4	-0.404	1.757

Observemos que la variable cuya F asociada es mayor es justamente la que acabamos de sacar, es decir, X_4 . Por lo tanto, terminamos el proceso.

En conclusión, según el método de Inclusión Progresiva, el mejor modelo es: $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i$. El modelo obtenido es consistente con lo que esperábamos luego de observar la matriz de correlación entre las variables.

Para este problema, el método de Eliminación Progresiva nos da el mismo modelo como el mejor. Esto no es cierto en general: cada método puede producir un resultado distinto.

6.5 Consideraciones importantes

No puede olvidarse nunca el importantísimo papel que juegan el sentido común y el conocimiento del problema en el análisis estadístico del mismo. Los procedimientos automáticos para selección de modelos pueden ser una herramienta útil, pero no deben ser usados en forma indiscriminada, pues en ocasiones producen resultados carentes de sentido a la luz del problema particular que se está estudiando.

Así mismo, no debe olvidarse nunca que un modelo lineal se basa en ciertas suposiciones que deben ser comprobadas. El que un modelo haya sido seleccionado como “el mejor” por algún procedimiento automático no garantiza que estas suposiciones se cumplan. Es siempre necesario, por tanto, analizar los residuos del modelo obtenido, y en caso de que se presenten problemas éstos deben ser resueltos.