

**Ejercicio** Se ajustó un modelo de regresión lineal a un conjunto de 40 observaciones que corresponden a un estudio sobre el aprovechamiento de la tierra y los desperdicios industriales. Las variables consideradas son:

- $Y$  = desperdicios sólidos (en millones de toneladas)
- $X_1$  = cantidad de tierra con fines industriales (en acres)
- $X_2$  = cantidad de tierra destinada a fabricaciones en metales (en acres)
- $X_3$  = cantidad de tierra destinada a caminos (en acres)

Se ajustó el modelo

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \epsilon_i \quad i = 1, \dots, 40$$

En la **Salida 1** se presentan los resultados obtenidos

- a) ¿Qué puede decir a partir de la tabla de correlaciones entre las variables? Si tuviera que ajustar un modelo con una sola variable, ¿cuál elegiría?
- b) A partir de la tabla de coeficientes estimados, ¿qué variables son significativas. ¿A qué nivel? ¿Cuál es el valor de  $s$ ? Especifique claramente las hipótesis nula y alternativa de cada uno de los tests de  $t$  reportados en dicha tabla. ¿Cómo se calcula el valor  $p$  de estos tests?
- c) ¿Es la regresión significativa? Especifique claramente las hipótesis nula y alternativa de este test. ¿Cómo se calcula el valor  $p$  en este caso? ¿Rechazaría al nivel 0.05?
- d) ¿Cómo se calcula la matriz de correlación de los estimadores de los coeficientes reportada en la Salida 1 ?
- e) Se aplicó un tests de  $F$  para testear la hipótesis

$$H_0 : \beta_1 = \beta_2 = 0$$

obteniéndose el siguiente resultado:

F	glnum	glden	p-valor
1.649439	2	36	0.206349

Interprete los resultados obtenidos e indique la decisión que toma. ¿Cómo se compara su decisión con la conclusión a la que arribó a partir del ítem b)?

- f) ¿Qué supuestos necesita realizar para que las conclusiones anteriores sean válidas?

Salida 1  
Matriz de Correlacion

	Y	X1	X2	X3
Y	1.0000000	0.1786278	0.4832167	0.6434090
X1	0.1786278	1.0000000	0.3909277	0.4186096
X2	0.4832167	0.3909277	1.0000000	0.8926119
X3	0.6434090	0.4186096	0.8926119	1.0000000

summary(salida)

Call: lm(formula = Y ~ X1 + X2 + X3)

Residuals:

Min	1Q	Median	3Q	Max
-0.5698	-0.1391	-0.02265	0.07987	0.8265

Coefficients:

	Value	Std. Error	t value	Pr(> t )
(Intercept)	0.2848	0.0484	5.8851	0.0000
X1	0.0000	0.0000	-0.7488	0.4588
X2	-0.0004	0.0003	-1.6217	0.1136
X3	0.0005	0.0001	3.9234	0.0004

Residual standard error: 0.2756 on 36 degrees of freedom

Multiple R-Squared: 0.4632

F-statistic: 10.35 on 3 and 36 degrees of freedom, the p-value is 0.00004695

Correlation of Coefficients:

	(Intercept)	X1	X2
X1	-0.1656		
X2	-0.0074	-0.0422	
X3	-0.1381	-0.1679	-0.8721