

Ejercicio # 3

Ronald Bailey

2023-05-28

Ejercicio #3:

A continuación se le muestran tres imágenes que muestran los resultados obtenidos de correr la función `summary()` a dos modelos de regresión lineal, para este ejercicio se le solicita que realice la interpretación de las tablas resultantes. Recuerde tomar en cuenta la significancia de los parámetros (significancia local), la significancia del modelo (significancia global), el valor del r cuadrado: y cualquier observación que considere relevante para determinar si el modelo estructuralmente es adecuado o no.

Modelo #1

La fórmula del modelo es $ROLL \sim UNEM$, lo que indica que la variable de respuesta es “ROLL” y la variable predictora es “UNEM”.

Los coeficientes del modelo son los siguientes:

- **Intercepto (Intercept):** El valor del intercepto es 3957. Esto significa que cuando la variable UNEM es cero, se espera que la variable de respuesta ROLL sea 3957. El valor p asociado con este coeficiente es 0.3313, que es mayor que 0.05, lo que significa que este coeficiente no es estadísticamente significativo al nivel del 5%. En otras palabras, no hay evidencia suficiente para rechazar la hipótesis nula de que este coeficiente sea cero.
- **UNEM:** El coeficiente para UNEM es 1133.8, lo que significa que por cada incremento unitario en UNEM, se espera que ROLL aumente en 1133.8 unidades, asumiendo que todas las demás variables se mantienen constantes. El valor p asociado con este coeficiente es 0.0358, que es menor que 0.05, indicando que este coeficiente es estadísticamente significativo al nivel del 5%.
- **El modelo tiene un R -cuadrado de 0.1531**, lo que indica que el modelo explica el 15.31% de la variabilidad en la variable de respuesta ROLL. El R -cuadrado ajustado es 0.1218, que toma en cuenta el número de predictores en el modelo (en este caso, solo uno).
- **El valor F es 4.883 con un valor p de 0.03579.** Esto indica que el modelo en su conjunto es estadísticamente significativo al nivel del 5%, es decir, hay evidencia de que al menos uno de los coeficientes del modelo es distinto de cero.

En resumen, aunque el modelo es significativo, el R -cuadrado es bastante bajo, lo que indica que el modelo no explica una gran cantidad de la variabilidad en los datos. Además, solo la variable predictora UNEM es significativa, mientras que el intercepto no es significativo. Esto podría sugerir que el modelo puede no ser el mejor ajuste para estos datos y que podrían ser necesarios más predictores o un modelo más complejo.

Modelo #2

- Intercepto (Intercept): El intercepto del modelo es $-9.153e+03$. Representa el valor esperado de la variable de respuesta ROLL cuando todas las variables predictoras son cero. Dado que el valor p asociado es muy pequeño ($5.02e-09$, mucho menor que 0.05), este coeficiente es estadísticamente significativo.
- UNEM: El coeficiente de UNEM es $4.501e+02$, lo que indica que por cada incremento unitario en UNEM, se espera que ROLL aumente en $4.501e+02$ unidades, manteniendo constantes todas las demás variables. Dado que el valor p asociado es muy pequeño (0.000807 , mucho menor que 0.05), este coeficiente es estadísticamente significativo.
- HGRAD: El coeficiente de HGRAD es $4.065e+01$, lo que indica que por cada incremento unitario en HGRAD, se espera que ROLL aumente en $4.065e+01$ unidades, manteniendo constantes todas las demás variables. Dado que el valor p asociado es muy pequeño ($1.52e-05$, mucho menor que 0.05), este coeficiente es estadísticamente significativo.
- INC: El coeficiente de INC es 4.275, lo que indica que por cada incremento unitario en INC, se espera que ROLL aumente en 4.275 unidades, manteniendo constantes todas las demás variables. Dado que el valor p asociado es muy pequeño ($5.59e-9$, mucho menor que 0.05), este coeficiente es estadísticamente significativo.
- El modelo tiene un R-cuadrado de 0.9621, lo que indica que el modelo explica el 96.21% de la variabilidad en la variable de respuesta ROLL. El R-cuadrado ajustado, que toma en cuenta el número de predictores en el modelo, es 0.9576.
- El valor F es 211.5 con un valor p muy pequeño ($< 2.2e-16$), lo que indica que el modelo en su conjunto es estadísticamente significativo.

En resumen, este modelo parece ser un buen ajuste para los datos, ya que tiene un R-cuadrado alto y todos los coeficientes son estadísticamente significativos. Sin embargo, siempre es importante tener en cuenta las suposiciones de la regresión lineal cuando se interpreta la salida del modelo.

Modelo #3

- Intercepto (Intercept): El valor del intercepto es 72.6826. Esto significa que cuando la variable Months es cero, se espera que la variable de respuesta Cab.Price sea 72.6826. El valor p asociado con este coeficiente es muy pequeño ($6.92e-10$, mucho menor que 0.05), lo que indica que este coeficiente es estadísticamente significativo.
- Months: El coeficiente para Months es 4.8626. Esto significa que por cada incremento unitario en Months, se espera que Cab.Price aumente en 4.8626 unidades, asumiendo que todas las demás variables se mantienen constantes. El valor p asociado con este coeficiente es muy pequeño ($7.18e-08$, mucho menor que 0.05), lo que indica que este coeficiente es estadísticamente significativo.
- El modelo tiene un R-cuadrado de 0.9509, lo que indica que el modelo explica el 95.09% de la variabilidad en la variable de respuesta Cab.Price. El R-cuadrado ajustado, que toma en cuenta el número de predictores en el modelo, es 0.946.
- El valor F es 193.6 con un valor p muy pequeño ($7.181e-08$), lo que indica que el modelo en su conjunto es estadísticamente significativo.

En resumen, este modelo de regresión parece ser un buen ajuste para los datos, dado que el R-cuadrado es alto y todos los coeficientes son estadísticamente significativos. No obstante, siempre es importante tener en cuenta las suposiciones de la regresión lineal cuando se interpreta la salida del modelo.