1. Preguntas cortas:
2. El coeﬃciente de determinación R2 para una regresión lineal múltiple con n = 10 observaciones y p = 5 predictores es igual a R2 = 0,86 y el valor F es igual a F = 20,5 ¿Cuál es el valor del coeﬃciente de determinación ajustado adjR2?

No se puede aun

(b) Decida (con breves explicaciones) si las siguientes afirmaciones son verdaderas o falsas.

1. Consideramos el modelo . Sea el intervalo de confianza del 95% para . En este caso, una prueba t con nivel de signiﬁcancia del 1% rechaza la hipótesis nula

RTA: Si

1. Los modelos complicados con muchos parámetros son mejores para la predicción que los modelos simples con pocos parámetros.

RTA: No necesariamente porque se debe establecer la relación con los demás predictores

1. En un modelo de regresión un factor con 4 niveles sirvió como predictor. Uno de estos niveles no es significativo y, en consecuencia, debe eliminarse del análisis.

No se puede aun

1. En una regresión lineal múltiple hay dos predictores que no son significativas. La prueba F global es altamente significativa. Por esta razón, es mejor no eliminar ambos predictores simultáneamente.

RTA: No porque puede que al eliminar solo una variable el modelo puede ajustarse mejor y explicar la variabilidad de Y

1. Considere el siguiente gráfico de dispersión:

Gráfico, Gráfico de dispersión

Descripción generada automáticamente

Los símbolos diﬀerentes en el gráfico corresponden a los valores de dos grupos diﬀerentes. La variable de respuesta y la covariable son continuas, la variable indicadora codifica la respectiva pertenencia al grupo.

1. Las covariables x y z interactúan. Explique.

Si interactúa ya que a medida que aumenta x también lo hace y, cuando se incluye el factor de Z, se puede ver que el intercepto con Z=0 es positivo y cuando Z=1 es negativo

1. ¿Están correlacionadas x y z? Explique.

Si están correlacionadas porque cuando Z=0 los valores de X son menores a 1y cuando Z=1 los valores de X son mayores a 1

1. ¿Qué modelo se ajusta a estos datos? Escriba la ecuación del modelo. El resultado R es el siguiente:

Texto

Descripción generada automáticamente

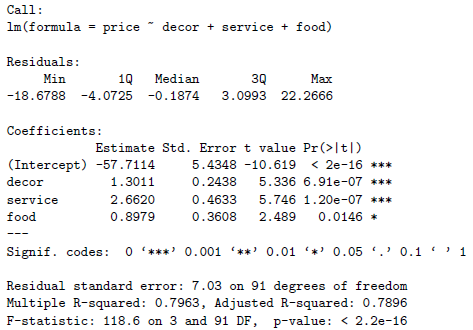
Y = 1.31 + 1.09x\_1 -1.25x\_2 -0.35XZ

1. ¿Cuáles son las líneas de regresión estimadas para los dos grupos?
2. Z=0; Y = 1.31 + 1.09x
3. Z=1; Y=(1.31-1.25)+(1.09+-035)x
4. ¿Es estadísticamente necesario ajustar dos rectas de regresión con pendientes diﬀerentes? Justifique su respuesta.
   1. En este caso si porque los datos e ajustan mejor individualmente por categoría
   2. Si se asume un nivel e significancia del 5% es mejor no tener dos rectas con pendientes diferentes
      1. Ho: B3 = 0
      2. Hi : B3 diferente 0
5. Esta pregunta tiene en cuenta la encuesta Zagat de 2008 sobre los restaurantes del barrio de River North de Chicago. Los datos contienen observaciones de 95 restaurantes que incluyen calificaciones de precio (en dólares), comida, decoración y servicio (en escalas discretas y ordenadas).
6. Comente la relación entre el precio y las agrupaciones creadas por las otras tres valoraciones. Sea conciso pero minucioso, señalando los patrones y los problemas.

Gráfico

Descripción generada automáticamente

1. Considere la salida resumida siguiente.



Describa brevemente el modelo que se está ajustando y cómo la(s) prueba(s) estadística(s) y otra información contenida en el resumen apoyan/refutan su(s) impresión(es) inicial(es) de la parte (a).