

ESTADÍSTICA

ESTUDIO DE REGRESIÓN Y CORRELACION SIMPLE

REGRESIÓN LINEAL MÚLTIPLE

Introducción

Es evidente que lo más económico y rápido para modelar el comportamiento de una variable Y es usar una sola variable preeditora y usar un modelo lineal. Pero algunas veces es bastante obvio de que el comportamiento de Y es imposible que sea explicada en gran medida por solo una variable.

Por ejemplo, es imposible tratar de explicar el rendimiento de un estudiante en un

examen, teniendo en cuenta solamente el número de horas que se preparó para ella. Claramente, el promedio académico del estudiante, la carga académica que lleva, el año de estudios, son tres de las muchas otras variables que pueden explicar su rendimiento. Tratar de explicar el comportamiento de Y con más de una variable preeditora usando una funcional lineal es el objetivo de regresión lineal múltiple.

Frecuentemente, uno no es muy familiar con las variables que están en juego y basa sus conclusiones solamente en cálculos obtenidos con los datos tomados.

Es decir, si ocurre que el coeficiente de determinación R^2 sale bajo (digamos menor de un 30%), considerando además que su valor no se ha visto afectado por datos anormales, entonces el modelo es pobre y para mejorarlo hay tres alternativas que frecuentemente se usan:

- a) Transformar la variable preeditora, o la variable de respuesta Y , o ambas y usar luego un modelo lineal.
- b) Usar regresión polinómica con una variable preeditora.
- c) Conseguir más variables preeditoras y usar una regresión lineal múltiple.

En el primer caso, se puede perder el tiempo tratando de encontrar la transformación más

adecuada y se podría caer en **overfitting**, es decir, encontrar un modelo demasiado optimista, que satisface demasiado la tendencia de los datos tomados pero que es pobre para hacer predicciones debido a que tiene una varianza grande.

En el segundo caso el ajuste es más rápido, pero es bien fácil caer en **overfitting** y, además se pueden crear muchos problemas de cálculo ya que pueden surgir problemas de colinealidad, es decir relación lineal entre los términos del modelo polinomio.

El tercer caso es tal vez la alternativa más usada y conveniente. Tiene bastante analogía con el caso simple, pero requiere el uso de vectores y matrices.

En el siguiente ejemplo se mostrará el uso interactivo de las tres alternativas a través de seis

favorezca en la situación aplicada, ya sea para tomar decisiones ajustadas con respecto al problema.

METODOLOGÍA.

La información fue recabada por medio de la empresa Grupo Ortega, de acuerdo a los registros archivados, respecto a los fletes realizados en el año 2001. Sobre los fletes más distantes de su origen para evaluar perfectamente la información recabada.

Posteriormente se efectuó el estudio de regresión múltiple ayudándonos con el programa de statgraphic.

PLANTEAMIENTO.

En la mediana empresa con el nombre Grupo Ortega, dedicada al transporte de todo tipo con destinos a toda la República Mexicana, ubicada en la carretera Guadalajara Colima. Km. 3.49 en Ciudad Guzmán Jalisco, se realizó un estudio para saber si influyen los gastos en el transcurso del viaje, con la liquidación del flete.

Los gastos a evaluar son:

- Combustible.
- Autopistas.
- Trabajos extramantenimiento.
- Viáticos (Comidas, transporte u otros).
- Llamadas a oficina.
- Mantenimiento

En el cual la variable independiente es el gasto de Liquidación.

CONCLUSIÓN

Con el estudio realizado podemos concluir que la relación de los gastos efectuados en los viajes arroja como resultado de relación .63.23%, lo cual no es buena relación.

Con el análisis de comparativo nos dimos cuenta de que las variables afectan en la ecuación de regresión pero no significativamente, la variable que más afecta en la ecuación de regresión fue la del combustible ya que creemos que por que fue en mayor cantidad pero la influencia que tiene es mínima ya que la liquidación total no varía mucho en los cambios de dicha variable.